

Dokumentvorlage, Version vom 20.01.2011

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

Fingolimod (Gilenya[®])

Novartis Pharma GmbH

Modul 4 A

Schubförmig remittierende Multiple Sklerose

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Stand: 21.09.2011

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	4
Abbildungsverzeichnis	9
Abkürzungsverzeichnis.....	11
4.1 Zusammenfassung der Inhalte von Modul 4.....	14
4.2 Methodik.....	36
4.2.1 Fragestellung	36
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung	38
4.2.3 Informationsbeschaffung.....	42
4.2.3.1 Studien des pharmazeutischen Unternehmers	42
4.2.3.2 Bibliografische Literaturrecherche	42
4.2.3.3 Suche in Studienregistern	44
4.2.3.4 Selektion relevanter Studien	45
4.2.4 Bewertung der Aussagekraft der Nachweise.....	45
4.2.5 Informationssynthese und -analyse	47
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien	47
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	47
4.2.5.3 Meta-Analysen.....	53
4.2.5.4 Sensitivitätsanalysen.....	54
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	56
4.2.5.6 Indirekte Vergleiche	57
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen	62
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	62
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel	62
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	62
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	64
4.3.1.1.3 Studien aus der Suche in Studienregistern	68
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	69
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	70
4.3.1.2.1 Studiendesign und Studienpopulationen	70
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	76
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	78
4.3.1.3.1 Endpunkte – RCT.....	78
4.3.1.3.1.1 Schübe	80
4.3.1.3.1.2 Behinderungsprogression	91
4.3.1.3.1.3 Gesundheitsbezogene Lebensqualität.....	99
4.3.1.3.1.4 Unerwünschte Ereignisse	104
4.3.1.3.2 Subgruppenanalysen – RCT.....	115

4.3.1.3.3	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	146
4.3.2	Weitere Unterlagen.....	155
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien.....	155
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	155
4.3.2.1.1.1	Studien des pharmazeutischen Unternehmers	156
4.3.2.1.1.2	Studien aus der bibliografischen Literaturrecherche	158
4.3.2.1.1.3	Studien aus der Suche in Studienregistern	166
4.3.2.1.1.4	Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel	169
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	172
4.3.2.1.2.1	Studiendesign und Studienpopulationen	172
4.3.2.1.2.2	Verzerrungspotenzial auf Studienebene	182
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	185
4.3.2.1.3.1	Endpunkte – indirekte Vergleiche aus RCT	185
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	213
4.3.2.2	Nicht randomisierte vergleichende Studien.....	214
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	214
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	214
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	215
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien....	215
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	216
4.3.2.3	Weitere Untersuchungen.....	217
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	217
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	217
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen.....	217
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	217
4.3.2.3.3.2	Subgruppenanalysen – Weitere Untersuchungen	218
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen.....	218
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	220
4.4.1	Beurteilung der Aussagekraft der Nachweise	220
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	221
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	224
4.4.4	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens....	224
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	225
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	225
4.5.2	Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	226
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	226
4.5.4	Verwendung von Surrogatendpunkten.....	226
4.6	Liste der eingeschlossenen Studien.....	228

4.7 Referenzliste.....	229
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche	237
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....	253
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund	256
Anhang 4-D : Liste der abgebrochenen Studien	257
Anhang 4-E : Liste der laufenden Studien.....	258
Anhang 4-F : Methodik der eingeschlossenen Studien – RCT.....	259
Anhang 4-G : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten	345
Anhang 4-H : Ergebnistabellen Subgruppenanalysen.....	602
Anhang 4-I : Kaplan-Meier Darstellungen.....	794

Tabellenverzeichnis

	Seite
Tabelle 4-1: Bewertung des Zusatznutzens innerhalb der relevanten Patientenpopulationen	34
Tabelle 4-2: Ein- und Ausschlusskriterien für Studien zur Bewertung des Nutzens und Zusatznutzens	38
Tabelle 4-3: Bewertung des Modells der Netzwerkanalyse.....	61
Tabelle 4-4: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	63
Tabelle 4-5: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	64
Tabelle 4-6: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	68
Tabelle 4-7: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	69
Tabelle 4-8: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	71
Tabelle 4-9: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	72
Tabelle 4-10: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	73
Tabelle 4-11: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	77
Tabelle 4-12: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	78
Tabelle 4-13: Operationalisierung von „Schübe“ ^{aa}	80
Tabelle 4-14: Bewertung des Verzerrungspotenzials für „Schübe“ in RCT mit dem zu bewertenden Arzneimittel	81
Tabelle 4-15: Ergebnisse für „Jährliche Schubrate“ aus RCT mit dem zu bewertenden Arzneimittel.....	82
Tabelle 4-16: Ergebnisse für „Anteil schubfreier Patienten“ aus RCT mit dem zu bewertenden Arzneimittel	84
Tabelle 4-17: Ergebnisse für „Schubcharakteristika“ aus RCT mit dem zu bewertenden Arzneimittel.....	87
Tabelle 4-18: Operationalisierung von „Behinderungsprogression“	91
Tabelle 4-19: Bewertung des Verzerrungspotenzials für „Behinderungsprogression“ in RCT mit dem zu bewertenden Arzneimittel	92
Tabelle 4-20: Ergebnisse für „Anteil Patienten ohne Behinderungsprogression“ aus RCT mit dem zu bewertenden Arzneimittel	93

Tabelle 4-21: Ergebnisse für „Mittlere Änderung des EDSS gegenüber Baseline“ aus RCT mit dem zu bewertenden Arzneimittel	96
Tabelle 4-22: Ergebnisse für „Mittlere Änderung des MSFC gegenüber Baseline“ aus RCT mit dem zu bewertenden Arzneimittel	98
Tabelle 4-23: Operationalisierung von „gesundheitsbezogener Lebensqualität“	99
Tabelle 4-24: Bewertung des Verzerrungspotenzials für „gesundheitsbezogene Lebensqualität“ in RCT mit dem zu bewertenden Arzneimittel	100
Tabelle 4-25: Ergebnisse für „gesundheitsbezogene Lebensqualität“ aus RCT mit dem zu bewertenden Arzneimittel	101
Tabelle 4-26: Operationalisierung von „Unerwünschten Ereignissen“	104
Tabelle 4-27: Bewertung des Verzerrungspotenzials für „Unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel	105
Tabelle 4-28: Ergebnisse für „Unerwünschte Ereignisse“ aus RCT mit dem zu bewertenden Arzneimittel	106
Tabelle 4-29: Operationalisierung von „Schübe“ ^{aa}	116
Tabelle 4-30: Bewertung des Verzerrungspotenzials für „Schübe“ in RCT mit dem zu bewertenden Arzneimittel	117
Tabelle 4-31: Ergebnisse für „Jährliche Schubrate“ aus RCT mit dem zu bewertenden Arzneimittel (ITT).....	118
Tabelle 4-32: Ergebnisse für „Jährliche Schubrate“ aus RCT mit dem zu bewertenden Arzneimittel (IFN Non-Responder Def I).....	119
Tabelle 4-33: Ergebnisse für „Jährliche Schubrate“ aus RCT mit dem zu bewertenden Arzneimittel (IFN Non-Responder Def II).....	121
Tabelle 4-34: Ergebnisse für „Jährliche Schubrate“ aus RCT mit dem zu bewertenden Arzneimittel (Patienten mit schnell fortschreitender schwerer RRMS).....	123
Tabelle 4-35: Ergebnisse für „Anteil schubfreier Patienten“ aus RCT mit dem zu bewertenden Arzneimittel (ITT)	125
Tabelle 4-36: Ergebnisse für „Anteil schubfreier Patienten“ aus RCT mit dem zu bewertenden Arzneimittel (IFN Non-Responder Def I)	127
Tabelle 4-37: Ergebnisse für „Anteil schubfreier Patienten, “ aus RCT mit dem zu bewertenden Arzneimittel (IFN Non-Responder Def II)	129
Tabelle 4-38: Ergebnisse für „Anteil schubfreier Patienten“ aus RCT mit dem zu bewertenden Arzneimittel (Patienten mit schnell fortschreitender schwerer RRMS)	131
Tabelle 4-39: Operationalisierung von „Behinderungsprogression“ ^{aa}	133
Tabelle 4-40: Bewertung des Verzerrungspotenzials für „Behinderungsprogression“ in RCT mit dem zu bewertenden Arzneimittel	134
Tabelle 4-41: Ergebnisse für „Anteil Patienten ohne Behinderungsprogression“ aus RCT mit dem zu bewertenden Arzneimittel (ITT)	135
Tabelle 4-42: Ergebnisse für „Anteil Patienten ohne Behinderungsprogression“ aus RCT mit dem zu bewertenden Arzneimittel (IFN Non-Responder Def I)	137

Tabelle 4-43: Ergebnisse für „Anteil Patienten ohne Behinderungsprogression“ aus RCT mit dem zu bewertenden Arzneimittel (IFN Non-Responder Def II)	139
Tabelle 4-44: Ergebnisse für „Anteil Patienten ohne Behinderungsprogression“ aus RCT mit dem zu bewertenden Arzneimittel (Patienten mit schnell fortschreitender schwerer RRMS)	141
Tabelle 4-45: Ergebnisse für „Mittlere Änderung des EDSS bzw. MSFC“ aus RCT mit dem zu bewertenden Arzneimittel.....	143
Tabelle 4-46: Operationalisierung von „gesundheitsbezogene Lebensqualität“	144
Tabelle 4-47: Bewertung des Verzerrungspotenzials für „gesundheitsbezogene Lebensqualität“ in RCT mit dem zu bewertenden Arzneimittel	144
Tabelle 4-48: Ergebnisse für „Mittlere Änderung des EQ-5D (Index und VAS)“ aus RCT mit dem zu bewertenden Arzneimittel	145
Tabelle 4-49: Liste der Studien des pharmazeutischen Unternehmers – Studien für indirekte Vergleiche	157
Tabelle 4-50: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – Studien für indirekte Vergleiche.....	158
Tabelle 4-51: Relevante Studien aus der Suche in Studienregistern – Studien für indirekte Vergleiche	167
Tabelle 4-52: Studienpool – Studien für indirekte Vergleiche	170
Tabelle 4-53: Charakterisierung der eingeschlossenen Studien – Studien für indirekte Vergleiche	173
Tabelle 4-54: Charakterisierung der Interventionen – Studien für indirekte Vergleiche.....	176
Tabelle 4-55: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche	178
Tabelle 4-56: Verzerrungspotenzial auf Studienebene – Studien für indirekte Vergleiche... ..	183
Tabelle 4-57: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	185
Tabelle 4-58: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	186
Tabelle 4-59: Operationalisierung von „Jährliche Schubrate“	188
Tabelle 4-60: Bewertung des Verzerrungspotenzials für „Jährliche Schubrate“ in RCT für indirekte Vergleiche	192
Tabelle 4-61: Ergebnisse für „Jährliche Schubrate“ ^{ac} aus RCT für indirekte Vergleiche.....	193
Tabelle 4-62: Ergebnisse für „Jährliche Schubrate“ aus RCT für indirekte Vergleiche.....	194
Tabelle 4-63: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	195
Tabelle 4-64: Operationalisierung von „Anteil schubfreier Patienten“	197
Tabelle 4-65: Bewertung des Verzerrungspotenzials für „Anteil schubfreier Patienten“ in RCT für indirekte Vergleiche.....	200

Tabelle 4-66: Ergebnisse für „Anteil schubfreier Patienten ^{ac} “ aus RCT für indirekte Vergleiche	201
Tabelle 4-67: Ergebnisse für „Anteil schubfreier Patienten“ aus RCT für indirekte Vergleiche	202
Tabelle 4-68: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	203
Tabelle 4-69: Operationalisierung von „nach 3 Monaten bestätigte Behinderungsprogression“	205
Tabelle 4-70: Bewertung des Verzerrungspotenzials für “nach 3 Monaten bestätigte Behinderungsprogression“ in RCT für indirekte Vergleiche.....	206
Tabelle 4-71: Ergebnisse für “Risiko einer nach 3 Monaten bestätigten Behinderungsprogression ^{ac} “ aus RCT für indirekte Vergleiche.....	207
Tabelle 4-72: Ergebnisse für „Nach 3 Monaten bestätigte Behinderungsprogression“ aus RCT für indirekte Vergleiche.....	208
Tabelle 4-73: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	209
Tabelle 4-74: Operationalisierung von „Behandlungsabbrüche aufgrund von UE“	210
Tabelle 4-75: Bewertung des Verzerrungspotenzials für „Behandlungsabbrüche aufgrund von UE“ in RCT für indirekte Vergleiche	211
Tabelle 4-76: Ergebnisse für „Behandlungsabbrüche aufgrund von UE ^{ac} “ aus RCT für indirekte Vergleiche	212
Tabelle 4-77: Ergebnisse für „Behandlungsabbrüche aufgrund unerwünschte Ereignisse“ aus RCT für indirekte Vergleiche	213
Tabelle 4-78: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	215
Tabelle 4-79: Operationalisierung von <Endpunkt xxx>.....	215
Tabelle 4-80: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	216
Tabelle 4-81: Operationalisierung von <Endpunkt xxx>.....	218
Tabelle 4-82: Bewertung des Zusatznutzens innerhalb der relevanten Patientenpopulationen	221
Tabelle 4-83: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	224
Tabelle 4-84: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens	225
Tabelle 4-85 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel	257
Tabelle 4-86 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel.....	258
Tabelle 4-87 (Anhang): Studiendesign und -methodik für Studie FREEDOMS.....	260
Tabelle 4-88 (Anhang): Studiendesign und -methodik für Studie TRANSFORMS	271

Tabelle 4-89 (Anhang): Studiendesign und -methodik für Studie Bornstein et al., 1987.....	282
Tabelle 4-90 (Anhang): Studiendesign und -methodik für Studie BECOME	287
Tabelle 4-91 (Anhang): Studiendesign und -methodik für Studie INCOMIN.....	294
Tabelle 4-92 (Anhang): Studiendesign und -methodik für Studie Etemadifar et al., 2006....	300
Tabelle 4-93 (Anhang): Studiendesign und -methodik für Studie IFNB-MS-Study Group..	304
Tabelle 4-94 (Anhang): Studiendesign und -methodik für Studie MSCRG	310
Tabelle 4-95 (Anhang): Studiendesign und -methodik für Studie Johnson et al., 1995	315
Tabelle 4-96 (Anhang): Studiendesign und -methodik für Studie REGARD.....	320
Tabelle 4-97 (Anhang): Studiendesign und -methodik für Studie BEYOND	326
Tabelle 4-98 (Anhang): Studiendesign und -methodik für Studie EVIDENCE	333
Tabelle 4-99 (Anhang): Studiendesign und -methodik für Studie PRISMS	339
Tabelle 4-100 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie FREEDOMS.....	390

Abbildungsverzeichnis

	Seite
Abbildung 4-1: Jährliche Schubrate unter Fingolimod verglichen mit Placebo zu Monat 24 der Studie 2301 (FREEDOMS, ITT-Population).....	19
Abbildung 4-2: Jährliche Schubrate unter Fingolimod verglichen mit Placebo zu Monat 24 der Studie 2301 (FREEDOMS, IFN Non-Responder Def I)	19
Abbildung 4-3: Jährliche Schubrate unter Fingolimod verglichen mit Placebo zu Monat 24 der Studie 2301 (FREEDOMS, IFN Non-Responder Def II)	20
Abbildung 4-4: Jährliche Schubrate unter Fingolimod verglichen mit Placebo zu Monat 24 der Studie 2301 (FREEDOMS, Patienten mit schnell fortschr. Schwerer RRMS)	20
Abbildung 4-5: Jährliche Schubrate unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie zu Monat 12 der Studie 2302 (TRANSFORMS): ITT-Population.....	21
Abbildung 4-6: Jährliche Schubrate unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie zu Monat 12 der Studie 2302 (TRANSFORMS): IFN Non-Responder Def I.....	21
Abbildung 4-7: Jährliche Schubrate unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie zu Monat 12 der Studie 2302 (TRANSFORMS): IFN Non-Responder Def II	22
Abbildung 4-8: Jährliche Schubrate unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie zu Monat 12 der Studie 2302 (TRANSFORMS): Patienten mit schnell fortschr. Schwerer RRMS	22
Abbildung 4-9: Anteil schubfreier Patienten über die Zeit unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie IFN- β -1a i. m. in TRANSFORMS (ITT-Population)	23
Abbildung 4-10: Anzahl schwerer Schübe unter Fingolimod verglichen mit Placebo in FREEDOMS.....	25
Abbildung 4-11: Anzahl hospitalisierungspflichtiger Schübe unter Fingolimod verglichen mit Placebo in FREEDOMS	25
Abbildung 4-12: Anzahl schwerer Schübe unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie IFN- β -1a i. m. in TRANSFORMS.....	26
Abbildung 4-13: Anzahl hospitalisierungspflichtiger Schübe unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie IFN- β -1a i. m. in TRANSFORMS	26
Abbildung 4-14: Anteil Patienten ohne nach 3 Monaten bestätigten Behinderungsprogression in FREEDOMS (ITT-Population)	27
Abbildung 4-15: Änderung gegenüber Studienbeginn des EDSS-Werts und des MSFC z-Punktwerts in Monat 12 in TRANSFORMS (ITT-Population)	29
Abbildung 4-16: Flussdiagramm der bibliographischen Suche zu Fingolimod.....	66
Abbildung 4-17: Anteil schubfreier Patienten unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie IFN- β -1a i. m. in TRANSFORMS (ITT-Population)	86
Abbildung 4-18: Zeit bis zum Auftreten einer nach 3 Monaten bestätigten Behinderungsprogression in FREEDOMS (ITT-Population)	95

Abbildung 4-19: Unerwünschte Ereignisse bei $\geq 10\%$ der Patienten nach 24 Monaten in FREEDOMS (Sicherheitspopulation)	111
Abbildung 4-20: Unerwünschte Ereignisse bei $\geq 10\%$ der Patienten nach 12 Monaten in TRANSFORMS (Sicherheitspopulation).....	111
Abbildung 4-21: Flussdiagramm der bibliografischen Literaturrecherche zu IFN- β -Studien (aktiv-kontrolliert mit IFN- β -Präparaten; IFN- β -1a i. m./s. c. vs. IFN- β -1b s. c.) ...	161
Abbildung 4-22: Flussdiagramm der bibliografischen Literaturrecherche zu IFN- β -Studien (aktiv-kontrolliert mit IFN- β -Präparaten; IFN- β -1a i. m vs. IFN- β -1a s. c.)	162
Abbildung 4-23: Flussdiagramm der bibliografischen Literaturrecherche zu IFN- β -Studien (Placebo-kontrolliert)	163
Abbildung 4-24: Flussdiagramm der bibliografischen Literaturrecherche zu Glatirameracetat-Studien (aktiv-kontrolliert).....	164
Abbildung 4-25: Flussdiagramm der bibliografischen Literaturrecherche zu Glatirameracetat-Studien (Placebo-kontrolliert)	165
Abbildung 4-26: Netzwerkstruktur für Endpunkt „Jährliche Schubrate“	187
Abbildung 4-27: Netzwerkstruktur für Endpunkt „Anteil schubfreier Patienten“.....	196
Abbildung 4-28: Netzwerkstruktur für Endpunkt „Behinderungsprogression“	204
Abbildung 4-29: Netzwerkstruktur für Endpunkt „Behandlungsabbrüche aufgrund von UE“	209

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ALT	Alanin-Aminotransferase
ANCOVA	Analysis of Covariance
ARR	Annualized Relapse Rate (jährliche Schubrate)
CHMP	Committee for Medical Products for Human Use
CPMP	Committee for Proprietary Medicinal Products
CONSORT	Consolidated Standards of Reporting Trials
DIMDI	Deutsches Institut für Medizinische Dokumentation
DMT	Disease-modifying therapy (krankheitsmodifizierende Therapie)
DSMB	Data Safety Monitoring Board
DSS	Disability Status Scale
E	Efficacy (E1/E4/E9/E10 Richtlinie)
EDSS	Expanded Disability Status Scale
EEC	European Economic Community
EG	Europäische Gemeinschaft
EKG	Elektrokardiogramm
EMA	European Medicines Agency (Europäische Arzneimittelbehörde)
EQ-5D	European Quality of Life – 5 Dimensions
EWP	Efficacy Working Party
FDA	Food and Drug Administration
FTY720	Fingolimod
Gd	Gadolinium
HR	Hazard Ratio
i. m.	Intramuskulär
ICH	International Conference on Harmonization
ICTRP	International Clinical Trial Registry Platform
IFN- β	Interferon β
ITT	Intention to treat
k.A.	Keine Angabe
KI	Konfidenzintervall
M	Mean (Mittelwert)

mcg	Mikrogramm
MRT	Magnetresonanztomographie
MS	Multiple Sklerose
MSFC	Multiple Sclerosis Functional Composite
MTC	Mixed Treatment Comparison
n. d.	nicht durchgeführt
NIS	Nicht-interventionelle Studie
PASAT	Paced Auditory Serial Addition Test
PML	Progressive multifokale Leukenzephalopathie
PPMS	Primär progrediente Multiple Sklerose
RCT	Randomized Controlled Trial
RR	Relatives Risiko
RRMS	Relapsing Remitting Multiple Sclerosis (schubförmig remittierende Multiple Sklerose)
SD	Standard Deviation (Standardabweichung)
SE	Standard Error (Standardfehler)
SEM	Standard Error of the Mean (Standardfehler des Mittelwertes)
SGB	Sozialgesetzbuch
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegendes unerwünschtes Ereignis
T1/T2	Relaxationszeiten T1/T2
tgl.	täglich
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
UE	Unerwünschtes Ereignis
ULN	Upper limit of normal
VAS	Visual Analog Scale (visuelle Analogskala)
vs.	versus
WHO	World Health Organization
ZNS	Zentrales Nervensystem

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.

Für Arzneimittel, die zur Behandlung eines seltenen Leidens nach der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 über Arzneimittel für seltene Leiden zugelassen sind, müssen keine Nachweise zum medizinischen Nutzen und medizinischen Zusatznutzen vorgelegt werden (Abschnitt 4.3), solange der Umsatz des Arzneimittels mit der gesetzlichen Krankenversicherung zu Apothekenverkaufspreisen einschließlich Umsatzsteuer in den letzten 12 Monaten einen Betrag von 50 Millionen Euro nicht übersteigt. Angaben zu Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, müssen jedoch vorgelegt werden. Zu diesem Zweck enthält die Dokumentvorlage den Abschnitt 4.4.4, der ausschließlich in Dossiers für die oben genannten Arzneimittel zur Behandlung eines seltenen Leidens bearbeitet werden soll. In diesem Abschnitt sind Angaben zu Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich der begründenden Nachweise vorzulegen. Weitere Abschnitte von Modul 4 können dabei nach Bedarf bearbeitet werden.

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

Es soll nachgewiesen werden, dass Fingolimod 0,5 mg einmal täglich oral bei Patienten mit hochaktiver RRMS hinsichtlich patientenrelevanter Endpunkte einen medizinischen Nutzen (Vergleich zu Placebo) beziehungsweise Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie (Vergleich zu IFN- β -1a i. m. 30 μ g einmal wöchentlich intramuskulär, beziehungsweise Glatirameracetat 20 mg einmal täglich subkutan) besitzt. Es soll dabei insbesondere geprüft werden, ob Patienten, die aufgrund mangelnder Therapietreue nicht auf IFN- β respondieren, durch Fingolimod einen Zusatznutzen gegenüber IFN- β -1a i. m. haben. Bei Patienten, die trotz ausreichender Therapietreue nicht auf IFN- β respondieren, soll ein Zusatznutzen gegenüber Glatirameracetat nachgewiesen werden. Der Zusatznutzen bei Patienten, die an einer schnell fortschreitenden schweren RRMS leiden, soll gegenüber IFN- β -1a i. m. nachgewiesen werden. Die Patientenpopulationen und die jeweiligen Vergleichstherapien sind durch ein Beratungsgespräch (1) zur zweckmäßigen Vergleichstherapie wie folgt definiert:

- Im Anwendungsgebiet zu a:
 - 1) Für Patienten, die nicht auf einen vollständigen und angemessenen normalerweise mindestens ein Jahr andauernden Zyklus einer Beta-Interferon-Therapie angesprochen haben: Glatirameracetat.
 - 2) Für Patienten, die noch keine ausreichende Therapie mit Beta-Interferonen erhalten haben: Beta-Interferon 1a oder 1b.
- Im Anwendungsgebiet zu b: Beta-Interferon 1a oder 1b

Die Unterteilung des Anwendungsgebiets in die Untergruppen a und b ergibt sich dabei aus der zugelassenen Indikation (s. Modul 2, Tabelle 2-3). Unter a1 sind somit therapietreue Patienten, die auf IFN- β nicht respondieren (Non-Responder im Anwendungsgebiet a1, compliant) zu verstehen, unter a2 nicht-therapietreue Patienten, die auf IFN- β nicht respondieren (Non-Responder im Anwendungsgebiet a2, non-compliant). Unter b sind Patienten mit rasch fortschreitender schwerer RRMS zu verstehen (Patienten mit rasch fortschreitender schwerer RRMS im Anwendungsgebiet b).

Die Nachweise sollen soweit verfügbar mittels randomisierter kontrollierter Studien (Evidenzstufe 1b) erfolgen. Sofern keine direkten Vergleichsstudien verfügbar sind, stützt sich der Nachweis auf indirekte Vergleiche auf Basis von RCT.

Datenquellen

Zur Bewertung der Fragestellung wurden Studien des pharmazeutischen Unternehmers herangezogen. Eine bibliographische Literaturrecherche war erforderlich, ergab aber für das zu bewertende Arzneimittel keine weiteren Studien. Für den zum Vergleich von Fingolimod und Glatirameracetat notwendigen indirekten Vergleich wurde eine bibliographische Literaturrecherche durchgeführt. Für den Netzwerkvergleich wurden insgesamt 13 relevante Studien identifiziert, inklusive der Fingolimod-Studien des pharmazeutischen Unternehmers.

Eine umfassende Suche in öffentlichen Studienregistern (www.clinicaltrials.gov; www.who.int/ictrp) ergab keine weiteren relevanten Studien zu Fingolimod oder Glatirameracetat.

Ein-/Ausschlusskriterien für Studien

In die Bewertung des medizinischen Nutzens und Zusatznutzens werden nur randomisierte kontrollierte Studien (RCT: Evidenzstufe 1b) eingeschlossen. Sofern RCT vorliegen, sind Studien niedrigeren Evidenzgrades für die Bewertung nicht weiter relevant. Ein Studienbericht oder eine Volltext-Publikation muss vorliegen.

Die für die Nutzen- und Zusatznutzenbewertung heranzuziehenden Studien müssen hinsichtlich Indikation und Patientenpopulation für die gemäß Zulassung in Frage kommenden Patienten aussagekräftig sein. Es sind daher nur Studien relevant, die RRMS-Patienten einschließen und dabei IFN Non-Responder und Patienten mit schnell fortschreitender schwerer RRMS nicht ausschließen. Die Ein- und Ausschlusskriterien in den einzelnen Studien müssen also mindestens einen Teil der in der Zulassung von Fingolimod genannten Patientengruppen abdecken. Gemäß dieser Kriterien konnten keine Studien für einen indirekten Vergleich von Glatirameracetat mit Fingolimod identifiziert werden, die die IFN Non-Responder definieren. Entweder war eine Vortherapie von vornherein ausgeschlossen, oder die Vortherapien konnten keine IFN beinhalten, weil die Studien vor Einführung der IFN-Präparate rekrutiert hatten. Ein indirekter Vergleich mit Bezug auf die für Fingolimod relevanten Populationen ist daher nicht durchführbar. Die Suche wurde für diese Fragestellung daher auf alle erwachsenen RRMS-Patienten erweitert. Die Studienpopulationen in den RCT sollen entsprechend der Zulassung für Fingolimod aus Erwachsenen (ab 18 Jahre) bestehen.

Für die Bewertung relevant sind Studien, die mindestens einen patientenrelevanten Endpunkt untersuchen (Morbidität, Mortalität und gesundheitsbezogene Lebensqualität). Studien, die ausschließlich Surrogatparameter als Endpunkte aufführen, sind nicht für die Bewertung relevant.

Nur Studien mit Fingolimod in der zugelassenen Dosierung von 0,5 mg sind relevant für den direkten Vergleich. Vergleichstherapien, die für die Studienauswahl relevant sind, sind Placebo um den medizinischen Nutzen von Fingolimod zu zeigen und IFN- β -1a i. m. sowie Glatirameracetat um den medizinische Zusatznutzen zu zeigen (siehe Begründung zur

zweckmäßigen Vergleichstherapie in Modul 3). Da die Recherche keine direkten Vergleichsstudien mit Glatirameracetat ergeben hat, wurde ein indirekter Vergleich notwendig. Für diese Netzwerkanalyse wurden Studien mit IFN- β -Präparaten und Glatirameracetat gewählt - mit den jeweils zugelassenen Dosierungen und Applikationsschemata und jeweils im Vergleich untereinander oder gegen Placebo.

Es wurden ausschließlich Studien mit einer Studiendauer ≥ 12 Monate in die Analyse eingeschlossen. Aufgrund der Tatsache, dass es sich um eine chronische Erkrankung handelt, wird eine kürzere Studiendauer als nur bedingt aussagekräftig gesehen. Eine Aussage zur Behinderungsprogression oder zur Entwicklung der Schubrate lässt sich in kürzeren Studien nicht mit ausreichender Sicherheit zur Beantwortung der obigen Fragestellung treffen. Insbesondere für den Endpunkt der jährlichen Schubrate ist eine kürzere Studiendauer nicht valide, da dann vorhandene Daten auf ein Jahr extrapoliert werden müssen, was zu einer unangemessenen Unschärfe der Aussage im Hinblick auf die im Dossier erörterte Fragestellung führt.

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

Das Verzerrungspotenzial der zur Bewertung herangezogenen Studien wurde anhand der endpunktübergreifenden und endpunktspezifischen Aspekte beurteilt. Details zum Verzerrungspotenzial sind den Bewertungsbögen in Anhang 4-G zu entnehmen.

Die Beschreibung des Designs und der Methodik der RCT erfolgte anhand der durch das CONSORT-Statement vorgegebenen Standards. Die Informationen zu den Items 2b bis 14 wurden jeweils dem Studienprotokoll und dessen Amendments entnommen (siehe Anhang 4-F).

In die Bewertung des medizinischen Nutzens und Zusatznutzens im direkten und indirekten Vergleich wurden ausschließlich RCT (Evidenzstufe 1b) eingeschlossen. Die Studien sind für die zu bewertende Indikation und die relevanten Patientengruppen aufgrund der Population (RRMS-Erkrankung, Krankheitsaktivität und Altersstruktur), Fallzahl, Studiendauer, Art der Intervention und der Wahl der Endpunkte relevant und erlauben eine valide Beurteilung des Nutzens und Zusatznutzens. Auch im Hinblick auf das Verzerrungspotenzial zeigen sich eine sehr hohe Studienqualität und Validität der herangezogenen Studien und deren Endpunkte.

Die Studien enthielten patientenrelevante Endpunkte zur Morbidität (Schübe, Behinderungsprogression, UE) und gesundheitsbezogenen Lebensqualität. Die Studiendauer der ausgewählten RCT beträgt mindestens 12 Monate. Die Dauer begründet sich mit den schubbezogenen und behinderungsprogressions-bezogenen Endpunkten, die erst ab ausreichender Studiendauer aussagekräftig sind. Um den medizinischen Nutzen von Fingolimod zu zeigen, wurde gegen Placebo getestet; der medizinische Zusatznutzen zeigte sich im Vergleich zur zweckmäßigen Vergleichstherapie IFN- β -1a i. m beziehungsweise Glatirameracetat.

Die Aussagekraft der oben genannten RCT ist somit als sehr hoch einzuschätzen und damit für eine valide Beurteilung des medizinischen Nutzens und Zusatznutzens uneingeschränkt geeignet. Die verwendeten Endpunkte sind außerdem, wie in Abschnitt 4.2.5.1 ausführlich erläutert, patientenrelevant.

Der indirekte Vergleich über die hier durchgeführte Netzwerkanalyse bietet für die Fragestellung des Zusatznutzens gegenüber Glatirameracetat die best-verfügbare Evidenz, da keine direkten Vergleichsstudien existieren. Die Methodik des indirekten Vergleichs folgt den Empfehlungen des G-BA und IQWiG und damit dem aktuellen Stand der Wissenschaft und wird daher als valide angesehen.

Die Ergebnisse der einzelnen Endpunkte der direkten Vergleiche werden für die Gesamtpopulation sowie für die relevanten Patientenpopulationen, wie in der Fragestellung formuliert, gegenübergestellt. Eine Meta-Analyse der Daten war nicht möglich, da nur eine Placebo-kontrollierte und eine aktiv-kontrollierte Studie vorliegen. Für den indirekten Vergleich wurden Endpunkte gewählt, die für alle relevanten Studien aus den jeweiligen Publikationen beziehungsweise Berichten auswertbar waren. Die Ergebnisse der Glatirameracetat-Studien wurden mittels Meta-Analyse zusammengefasst und dann im indirekten Vergleich (Netzwerkanalyse) mit Fingolimod dargestellt.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Der medizinische Nutzen und Zusatznutzen sowie die Sicherheit von Fingolimod wurde im bis dato größten Studienprogramm der MS-Forschung untersucht. Das Fingolimod-Studienprogramm (Phase II- und III-Studien) umfasst insgesamt mehr als 4.000 Patienten entsprechend einer Gesamtexposition von etwa 10.000 Patientenjahren, von denen mehr als 2.000 Patienten länger als 2 Jahre behandelt wurden. Ungefähr 130 dieser Patienten wurden bereits 6 Jahre oder länger behandelt (Therapie dauert derzeit noch an), so dass auch Langzeit-Sicherheitsdaten zur Verfügung stehen.

Zwei Studien erfüllten die oben genannten Kriterien für die Aufnahme in die Nutzen- und Zusatznutzenbewertung: die Placebo-kontrollierte 24-monatige Phase III-Studie FREEDOMS und die aktiv-kontrollierte (IFN- β -1a i. m.) 12-monatige Phase III-Studie TRANSFORMS. Für einen Vergleich von Fingolimod mit Glatirameracetat wurden keine direkten vergleichenden Studien gefunden. Die Zusatznutzenbewertung erfolgt daher über einen indirekten Vergleich.

Die in der Fragestellung dargestellten relevanten Patientenpopulationen wurden in der Auswertung der Studien möglichst genau berücksichtigt. Die Darstellung der Ergebnisse erfolgt für die Gesamtpopulation (ITT), die Gruppe der IFN Non-Responder und die Gruppe der Patienten mit schnell fortschreitender schwerer RRMS. Bei der Gruppe der IFN Non-Responder (Anwendungsgebiet a1 und a2) ist zu beachten, dass der Zulassungstext jeweils zwei Definitionen für einen Non-Responder zulässt (s. Modul 2, Tabelle 2-3), die als IFN Non-Responder Def I und IFN Non-Responder Def II berücksichtigt werden. Sowohl die IFN

Non-Responder Def I als auch die IFN Non-Responder Def II umfassen jeweils beide Anwendungsgebiete a1 und a2. Eine Unterscheidung nach therapietreuen (a1) und nicht therapietreuen (a2) Patienten war mit der verfügbaren Datengrundlage nicht möglich.

Schubbezogene Endpunkte

In beiden Phase III-Studien senkte Fingolimod 0,5 mg die jährliche Schubrate (ARR) signifikant verglichen mit den Vergleichsbehandlungen (Placebo und zweckmäßige Vergleichstherapie IFN- β -1a i. m.) und erfüllte erfolgreich die primären Ziele der Phase III-Hauptstudien. Fingolimod 0,5 mg senkte die jährliche Schubrate in FREEDOMS nach 2 Jahren der Behandlung signifikant um 54 % im Vergleich zu Placebo (ARR ratio 0,462; 95 %-KI (0,373; 0,571); $p < 0,001$; Tabelle 4-15; Abbildung 4-1). In TRANSFORMS senkte Fingolimod 0,5 mg die jährliche Schubrate nach 1 Jahr der Behandlung im Vergleich zu IFN- β -1a i. m. um 52 % (ARR ratio 0,484; 95 %-KI (0,367; 0,639); $p < 0,001$; Tabelle 4-15; Abbildung 4-5). In den Auswertungen zur Zielpopulation zeigt sich gegenüber Placebo in allen drei definierten Subpopulationen ein konsistenter signifikanter Effekt mit einer Senkung der jährlichen Schubrate zwischen 62 % und 71 % (ARR ratios und 95 %-KIs für IFN Non-Responder Def I und II und Patienten mit schnell fortschreitender schwerer RRMS: 0,379 (0,212; 0,677), 0,287 (0,158; 0,524), 0,329 (0,175; 0,620); Abbildung 4-2 bis Abbildung 4-4). Für die beiden Subpopulationen der IFN Non-Responder verhält es sich gegenüber der zweckmäßigen Vergleichstherapie IFN- β -1a i. m. ebenso. Die jährliche Schubrate wird dabei signifikant um 61 % gegenüber IFN- β -1a i. m. gesenkt. Die Senkung der jährlichen Schubrate um 25 % bei den Patienten mit schnell fortschreitender MS war hingegen nicht signifikant, die Gruppengröße mit rund 30 Patienten aber auch nicht aussagekräftig (ARR ratios und 95 %-KIs für IFN Non-Responder Def I und II und Patienten mit schnell fortschreitender schwerer RRMS: 0,391 (0,246; 0,623), 0,389 (0,244; 0,619), 0,746 (0,238; 2,333); Abbildung 4-6 bis Abbildung 4-8).

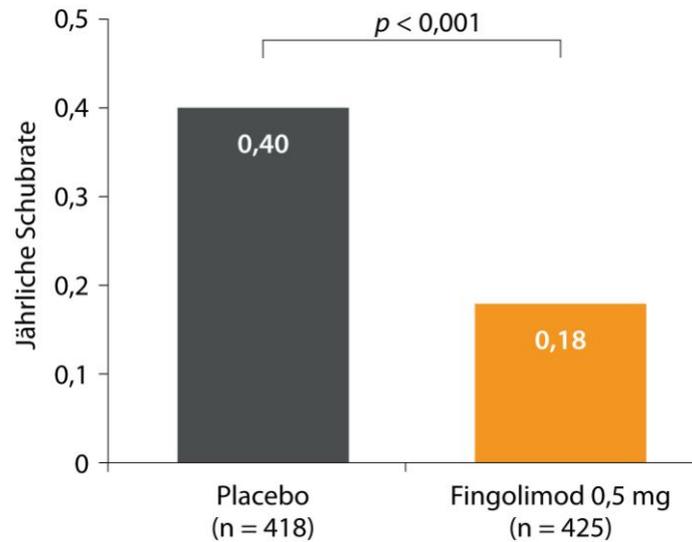


Abbildung 4-1: Jährliche Schubrate unter Fingolimod verglichen mit Placebo zu Monat 24 der Studie 2301 (FREEDOMS, ITT-Population)

Intent-to-treat Population. Der p -Wert für Fingolimod vs. Komparator wurde mit einem negativ binomialen Regressionsmodell errechnet und nach Behandlungsgruppe, Land, Anzahl der Schübe in den 2 Jahren vor Studienbeginn und Expanded Disability Status Scale (EDSS) Score bei Studienbeginn adjustiert.

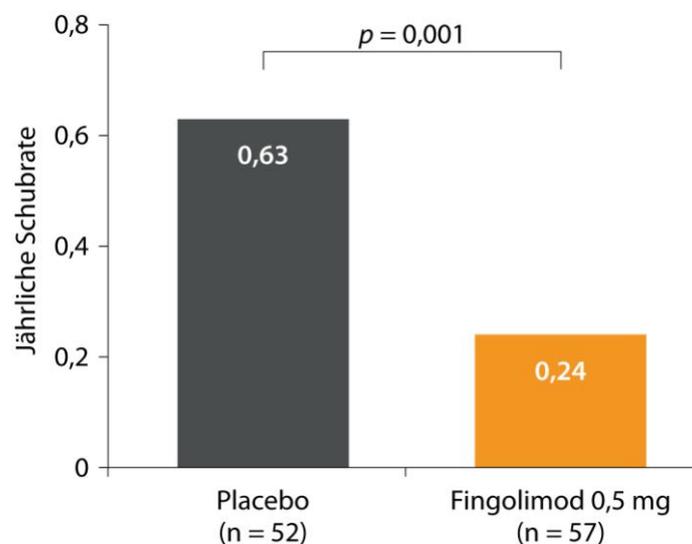


Abbildung 4-2: Jährliche Schubrate unter Fingolimod verglichen mit Placebo zu Monat 24 der Studie 2301 (FREEDOMS, IFN Non-Responder Def I)

Intent-to-treat Population. Der p -Wert für Fingolimod vs. Komparator wurde mit einem negativ binomialen Regressionsmodell errechnet und nach Behandlungsgruppe adjustiert.

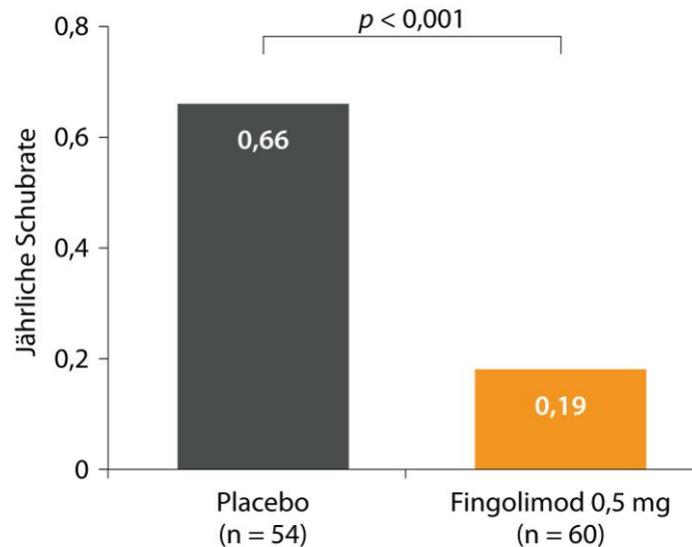


Abbildung 4-3: Jährliche Schubrate unter Fingolimod verglichen mit Placebo zu Monat 24 der Studie 2301 (FREEDOMS, IFN Non-Responder Def II)

Intent-to-treat Population. Der p -Wert für Fingolimod vs. Komparator wurde mit einem negativ binomialen Regressionsmodell errechnet und nach Behandlungsgruppe adjustiert.

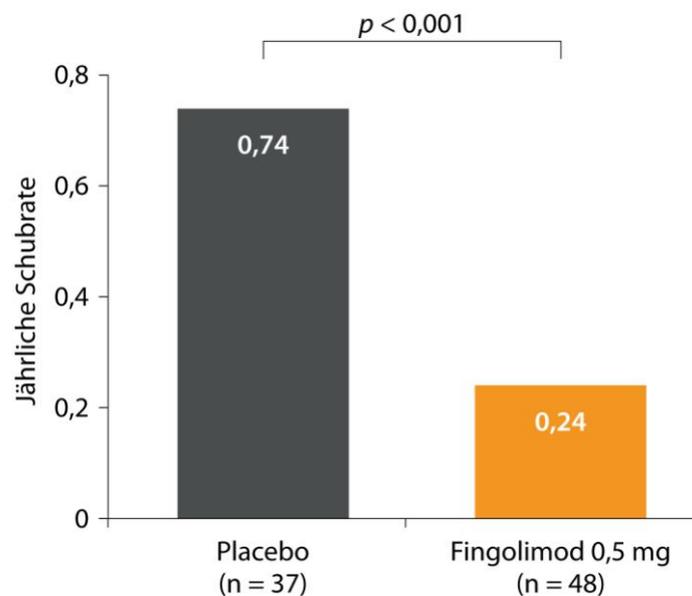


Abbildung 4-4: Jährliche Schubrate unter Fingolimod verglichen mit Placebo zu Monat 24 der Studie 2301 (FREEDOMS, Patienten mit schnell fortschr. Schwerer RRMS)

Intent-to-treat Population. Der p -Wert für Fingolimod vs. Komparator wurde mit einem negativ binomialen Regressionsmodell errechnet und nach Behandlungsgruppe adjustiert.

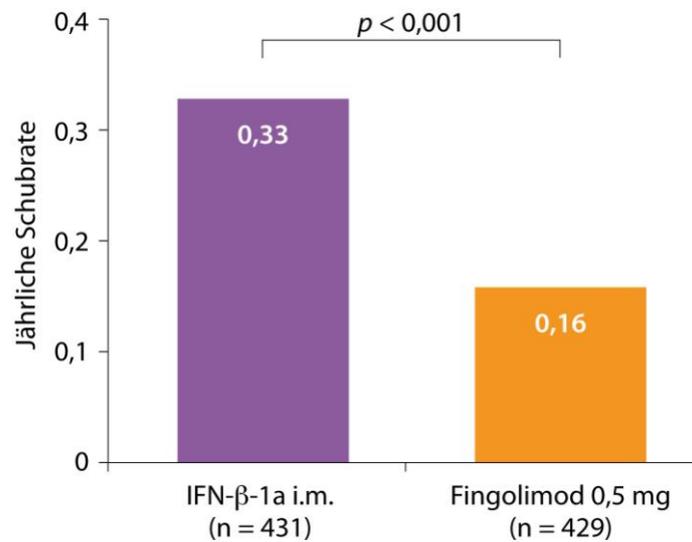


Abbildung 4-5: Jährliche Schubrate unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie zu Monat 12 der Studie 2302 (TRANSFORMS): ITT-Population

Intent-to-treat Population. Der p -Wert für Fingolimod vs. IFN- β -1a i. m. wurde mit einem negativ binomialen Regressionsmodell errechnet und nach Behandlungsgruppe, Land, Anzahl der Schübe in den 2 Jahren vor Studienbeginn und Expanded Disability Status Scale (EDSS) Score bei Studienbeginn adjustiert.

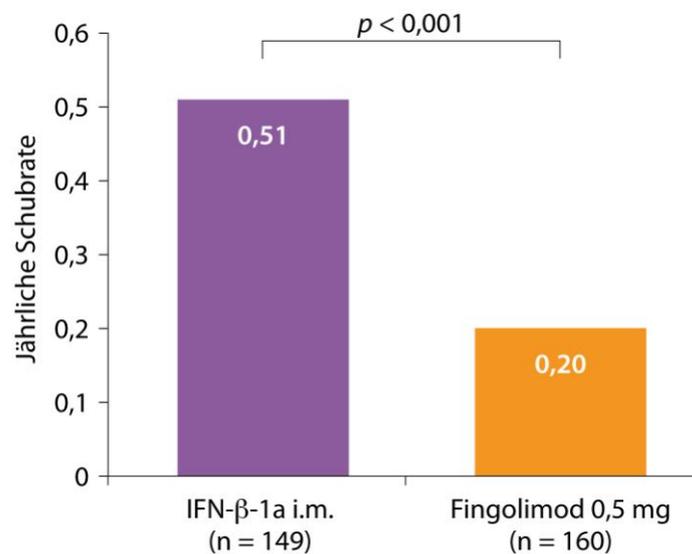


Abbildung 4-6: Jährliche Schubrate unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie zu Monat 12 der Studie 2302 (TRANSFORMS): IFN Non-Responder Def I

Intent-to-treat Population. Der p -Wert für Fingolimod vs. IFN- β -1a i. m. wurde mit einem negativ binomialen Regressionsmodell errechnet und nach Behandlungsgruppe adjustiert.

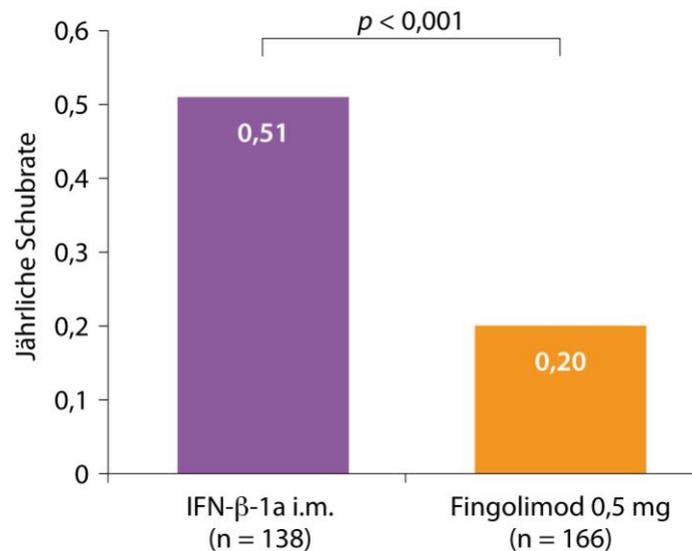


Abbildung 4-7: Jährliche Schubrate unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie zu Monat 12 der Studie 2302 (TRANSFORMS): IFN Non-Responder Def II

Intent-to-treat Population. Der p -Wert für Fingolimod vs. IFN-β-1a i. m. wurde mit einem negativ binomialen Regressionsmodell errechnet und nach Behandlungsgruppe adjustiert.

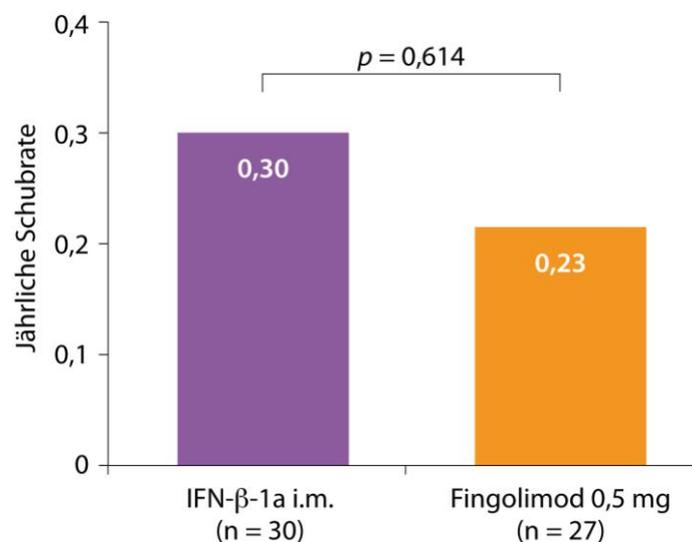


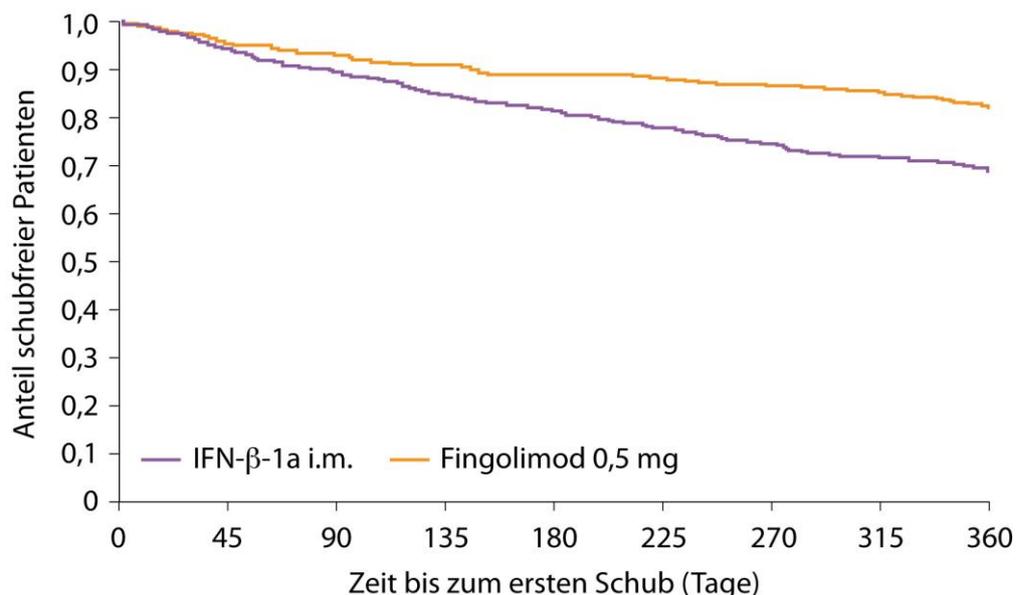
Abbildung 4-8: Jährliche Schubrate unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie zu Monat 12 der Studie 2302 (TRANSFORMS): Patienten mit schnell fortschr. Schwerer RRMS

Intent-to-treat Population. Der p -Wert für Fingolimod vs. IFN-β-1a i. m. wurde mit einem negativ binomialen Regressionsmodell errechnet und nach Behandlungsgruppe adjustiert.

Im indirekten Vergleich war die Schubrate unter Glatirameracetat 1,411-mal höher (95 %-KI: (1,125; 1,768)) als unter Fingolimod, umgekehrt ausgedrückt war die relative Schubrate unter Fingolimod 0,709. Die Schubrate war damit unter Fingolimod gegenüber Glatirameracetat um rund 30 % signifikant reduziert. Die durchgeführte Netzwerkanalyse kann für den

Wirksamkeitsvergleich zwischen Fingolimod und Glatirameracetat nur als Annäherung gelten und ist auch für die gemäß zugelassener Indikation in Frage kommenden Patienten nur mäßig aussagekräftig.

In beiden Phase III-Studien wurde der Anteil schubfreier Patienten in den Fingolimod-Behandlungsgruppen signifikant erhöht gegenüber den Vergleichsbehandlungen. In FREEDOMS waren 70 % (95 %-KI: (65,95; 74,80)) der mit Fingolimod 0,5 mg behandelten Patienten nach 24 Monaten schubfrei im Vergleich zu 46 % (95 %-KI: (40,70; 50,57)) unter Placebo. In TRANSFORMS waren 83 % (95 %-KI: (79,02; 86,27)) der mit Fingolimod 0,5 mg behandelten Patienten nach 12 Monaten schubfrei im Vergleich zu 69 % (95 %-KI: (64,84; 73,76)) der mit IFN- β -1a i. m. behandelten Patienten. Schon im Verlauf zeigte sich, dass unter Fingolimod 0,5 mg mehr Patienten schubfrei waren als unter IFN- β -1a i. m. (Abbildung 4-17). Zensierungen im Verlauf begründen sich auf Studienabbrüche (Darstellungen inkl. Zensierungen finden sich im Anhang 4-I).



Anzahl Patienten mit Risiko für den ersten Schub (number at risk)

Fingolimod 0,5 mg	429	408	397	385	374	368	359	351	294
IFN- β -1a i. m.	431	402	379	350	335	313	299	286	240

Abbildung 4-9: Anteil schubfreier Patienten über die Zeit unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie IFN- β -1a i. m. in TRANSFORMS (ITT-Population)

Intent-to-treat-Population. Daten werden als Prozentsatz von Patienten ohne bestätigten Schub basierend auf Kaplan-Meier-Schätzungen dargestellt.

In den der Zielpopulation entsprechenden Subpopulationen sind die Daten konsistent und zeigen ebenfalls einen signifikanten Vorteil gegenüber Placebo beziehungsweise gegenüber der zweckmäßigen Vergleichstherapie IFN- β -1a i. m., die Größe des Effekts spricht in den Subpopulationen (abgesehen von den Patienten mit schnell fortschreitender schwerer MS in

der TRANSFORMS-Studie) sogar noch deutlicher für Fingolimod im Vergleich zu den jeweiligen Komparatoren. Lediglich bei den Patienten mit schnell fortschreitender MS war der Effekt nicht signifikant unterschiedlich, die Gruppengröße mit rund 30 Patienten dafür aber auch nicht aussagekräftig. Kaplan-Meier-Darstellungen inklusive Zensierungen finden sich im Anhang 4-I, Zensierungen im Verlauf begründen sich auf Studienabbrüche.

Der Anteil schubfreier Patienten ist unter Fingolimod und Glatirameracetat vergleichbar. Zusammen mit der Aussage zur reduzierten Schubrate unter Fingolimod im Vergleich zu Glatirameracetat bedeutet das, dass zwar gleichviele Patienten Schübe haben, aber die Zahl der Schübe pro Patient unter Fingolimod im Vergleich zu Glatirameracetat reduziert wurde. Es gilt aber auch hier, dass die Interpretation der Daten mit Vorsicht erfolgen muss, da die Daten auf einem nur bedingt für die relevanten Patientenpopulationen aussagekräftigen indirekten Vergleich beruhen.

Insgesamt unterstützen die dargestellten Daten den klinischen Zusatznutzen von Fingolimod 0,5 mg gegenüber den zweckmäßigen Vergleichstherapien und zeigen eine deutliche Schubratenreduktion.

Die Gesamtzahl der Schübe und damit konsistent auch die Zahl schwerer Schübe oder Schübe, die einen Krankenhausaufenthalt beziehungsweise eine Steroidbehandlung erfordern, ist bei mit Fingolimod behandelten Patienten gegenüber mit IFN- β -1a i. m. behandelten Patienten reduziert (schwere Schübe 10 vs. 30, hospitalisierungspflichtige Schübe 11 vs. 36, steroidpflichtige Schübe 65 vs. 148; ITT Population). Diese Schübe sind im Hinblick auf die Kosten, die für das Gesundheitssystem entstehen, besonders relevant. Die Kosten, die durch einen hospitalisierungspflichtigen Schub entstehen, können bis um den Faktor 50 über jenen eines milden Schubes liegen, wenn die Kosten für Neuro-Rehabilitationsmaßnahmen einbezogen werden (2). Selbst ohne diese Aufwendungen können die Kosten einen Wert von ungefähr 2.300 EUR pro Schub erreichen, bei Notwendigkeit einer Plasmapherese sogar bis zu rund 9.000 EUR. Eine Reduktion der Anzahl hospitalisierungspflichtiger Schübe ist daher besonders bedeutsam. Die Reduktion der Anzahl schwerer Schübe und hospitalisierungspflichtiger Schübe war in der Gesamtpopulation sowie den gemäß Label relevanten Subpopulationen zu sehen. Besonders auffällig ist dabei, dass in der Subpopulation der Patienten mit schnell fortschreitender schwerer RRMS die Anzahl schwerer und hospitalisierungspflichtiger Schübe nicht nur konsistent mit der Gesamtzahl der Schübe sinkt, sondern der Anteil der Schübe dieser Schubkategorie sogar prozentual deutlich reduziert wird. Auch wenn also die Schubrate in dieser Subpopulation nur mäßig und nicht signifikant gesenkt werden kann, so zeigt sich doch, dass die Schubschwere deutlich abnimmt und damit ein reduzierender Effekt auf die dadurch entstehenden Kosten zu erwarten ist.

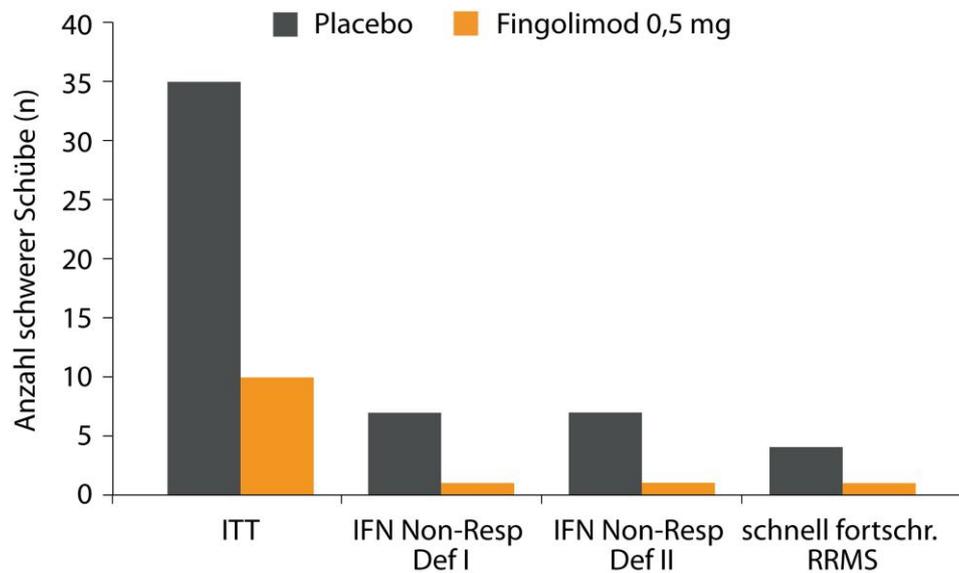


Abbildung 4-10: Anzahl schwerer Schübe unter Fingolimod verglichen mit Placebo in FREEDOMS

Intent-to-treat-Population. Daten werden als absolute Anzahl an Schüben dargestellt.

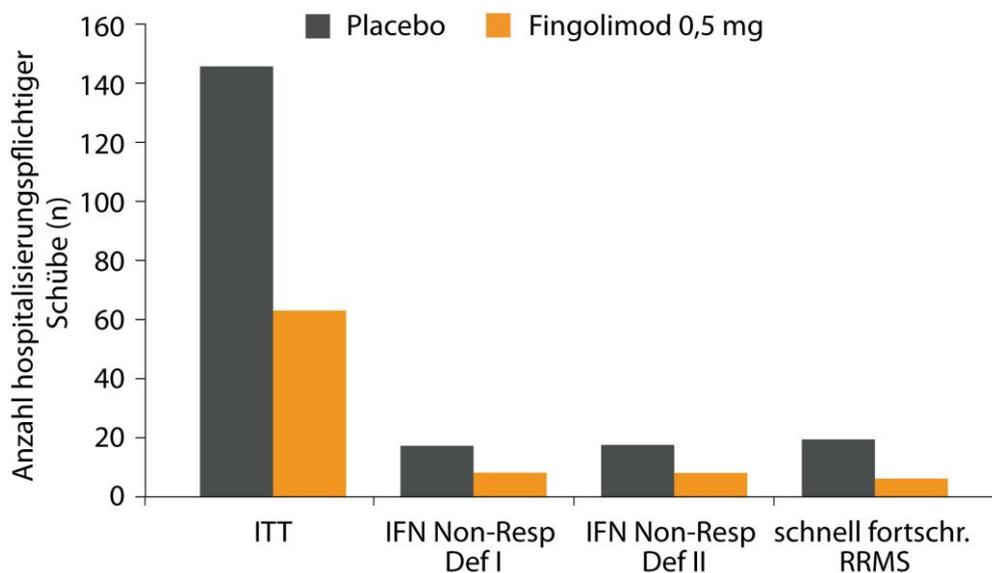


Abbildung 4-11: Anzahl hospitalisierungspflichtiger Schübe unter Fingolimod verglichen mit Placebo in FREEDOMS

Intent-to-treat-Population. Daten werden als absolute Anzahl an Schüben dargestellt.

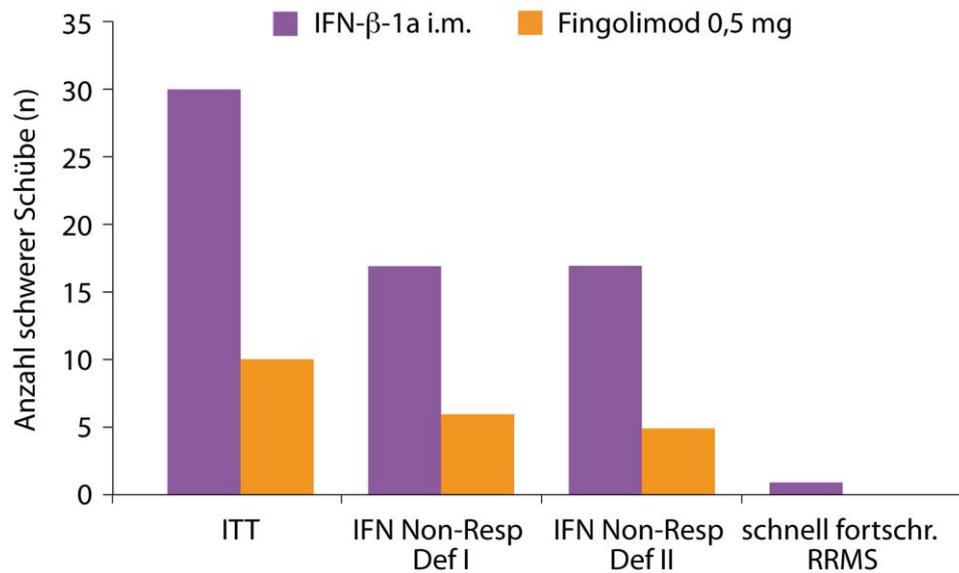


Abbildung 4-12: Anzahl schwerer Schübe unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie IFN-β-1a i. m. in TRANSFORMS Intent-to-treat-Population. Daten werden als absolute Anzahl an Schüben dargestellt.

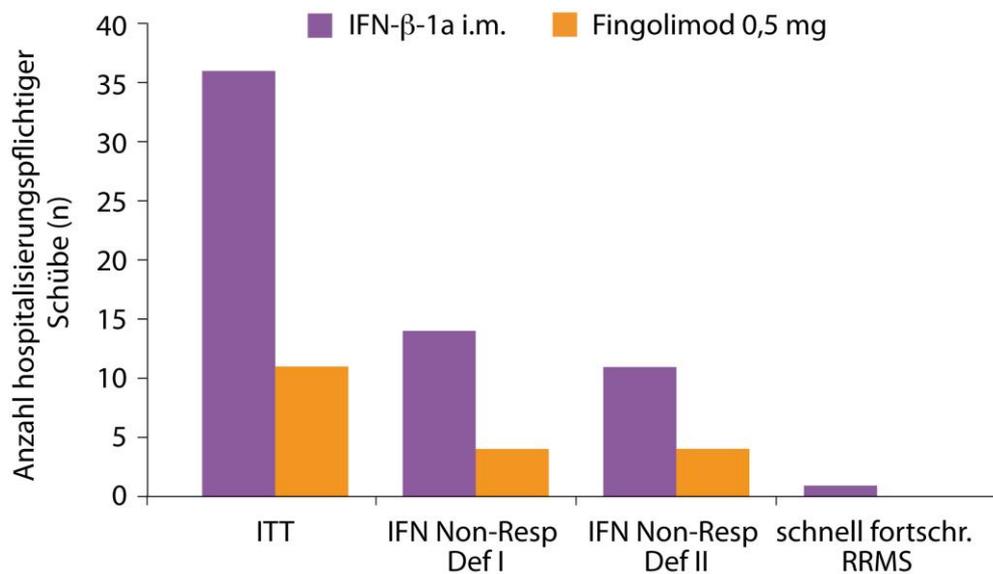
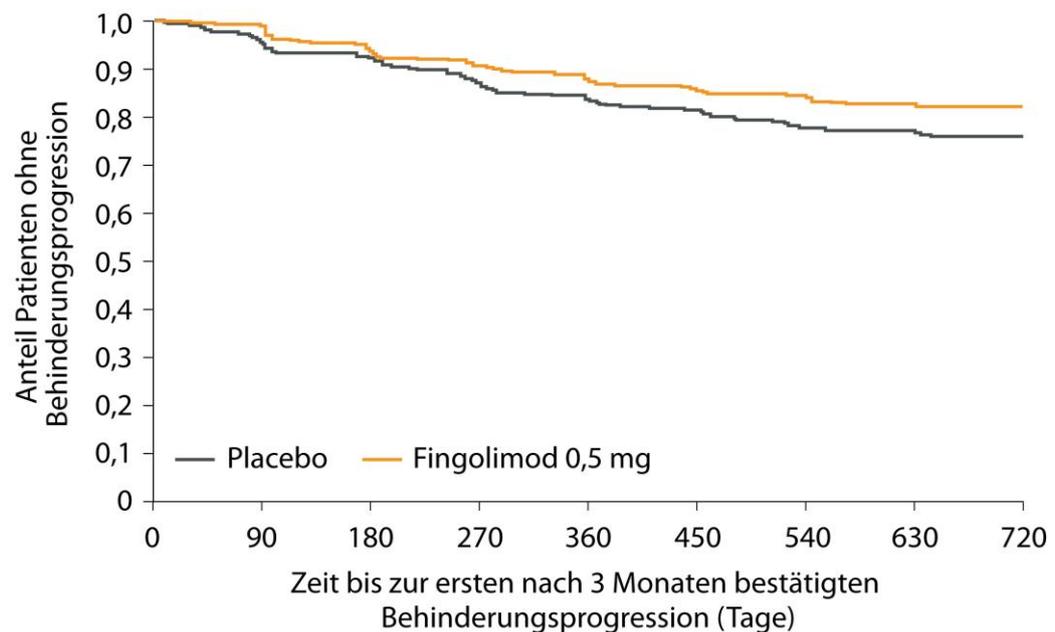


Abbildung 4-13: Anzahl hospitalisierungspflichtiger Schübe unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie IFN-β-1a i. m. in TRANSFORMS Intent-to-treat-Population. Daten werden als absolute Anzahl an Schüben dargestellt.

Endpunkte im Bezug auf Behinderungsprogression:

In FREEDOMS wurde das Risiko einer nach 3 Monaten bestätigten Behinderungsprogression über 2 Jahre unter Fingolimod 0,5 mg im Vergleich zu Placebo um 30 % reduziert (HR: 0,70; (95 %-KI: 0,52; 0,96), $p = 0,024$). Ein signifikant größerer Anteil der Patienten blieb unter Fingolimod 0,5 mg frei von einer bestätigten Progression der Behinderung als unter Placebo (siehe Abbildung 4-18). Zensierungen begründen sich auf Studienabbrechern, die hohe Zahl an Zensierungen in den letzten 3 Monaten vor allem auf nicht mehr bestätigte Fälle von Behinderungsprogression. Das sind Fälle, die nach Monat 21 auftraten, deren Termin für eine Bestätigung nach 3 Monaten aber nach Monat 24 gelegen hätten (Darstellungen inkl. Zensierungen finden sich in Anhang 4-I).



Patienten mit Risiko für die erste bestätigte Behinderungsprogression (number at risk)

Fingolimod 0,5 mg	425	416	388	370	354	340	332	321	152
Placebo	418	391	371	341	320	308	290	279	143

Abbildung 4-14: Anteil Patienten ohne nach 3 Monaten bestätigten Behinderungsprogression in FREEDOMS (ITT-Population)

Intent-to-treat-Population. Daten werden als Prozentsatz von Patienten mit bestätigter Behinderungsprogression basierend auf Kaplan-Meier-Schätzungen dargestellt. Die p -Werte für Fingolimod versus Placebo wurde mithilfe einer Log-Rank-Test-Hauptanalyse und eines proportionalen Risikomodells nach Cox berechnet, adjustiert nach Behandlung, Land, EDSS-Wert zu Studienbeginn und Alter.

Für die gemäß Zielpopulation definierten Subpopulationen war ein konsistenter Trend hin zu einem Vorteil von Fingolimod gegenüber Placebo zu sehen, die Ergebnisse waren jedoch nicht statistisch signifikant. Zusammengefasst deuten diese Daten darauf hin, dass Fingolimod übereinstimmende, klinisch bedeutsame Effekte auf die Behinderungsprogression hat. Kaplan-Meier-Darstellungen inklusive Zensierungen finden sich im Anhang 4-I, Zensierungen im Verlauf begründen sich auf Studienabbrüche.

Die Studie TRANSFORMS war aufgrund der Studiendauer von 12 Monaten nicht darauf ausgelegt, einen Unterschied in einer bestätigten Behinderungsprogression zwischen Fingolimod und IFN- β -1a i. m. nachzuweisen. Dennoch sprach die HR für eine bestätigte Progression nach 12 Monaten in TRANSFORMS für Fingolimod 0,5 mg (HR: 0,71; (95 %-KI: 0,42; 1,21), $p=0,209$). Die Trends der Wirkung scheinen mit den Erkenntnissen aus FREEDOMS übereinzustimmen. Dies gilt ebenso für die Subpopulationen gemäß Label.

Im indirekten Vergleich mit Glatirameracetat zeigte sich, dass die Rate der nach 3 Monaten bestätigten Behinderungsprogression unter Glatirameracetat und Fingolimod vergleichbar ist. Diese Daten beruhen, wie oben bereits erwähnt, nicht auf direkten Vergleichsstudien. Die durchgeführte Netzwerkanalyse kann nur als Annäherung gelten und ist auch für die gemäß zugelassener Indikation in Frage kommenden Patienten nur mäßig aussagekräftig.

Die Daten der Phase III-Studien zeigen, dass die mittleren EDSS-Werte und MSFC z-Punktwerte unter Fingolimod-Behandlung stabil blieben oder sich leicht verbesserten, während sie sich in den Gruppen (ITT-Population) mit Placebo und IFN- β -1a i. m. jedoch leicht verschlechterten. Sowohl in FREEDOMS als auch in TRANSFORMS sprachen die EDSS-Werte und MSFC z-Punktwerte für Fingolimod 0,5 mg im Vergleich zu den jeweiligen Vergleichsbehandlungen.

In den Subpopulationen zeigt sich ein ähnliches Bild. Tendenziell bietet Fingolimod hier einen Vorteil gegenüber den Vergleichsbehandlungen. Bei den IFN Non-Respondern ist der Vorteil von Fingolimod gegenüber IFN- β -1a i. m. für den MSFC statistisch signifikant:

IFN Non-Responder Def I: Mittlere (\pm SEM) Veränderung des EDSS $-0,09 \pm 0,063$ vs. $-0,04 \pm 0,073$; $p=0,116$ und des MSFC $0,02 \pm 0,028$ vs. $-0,13 \pm 0,057$; $p=0,006$;

IFN non-repsonder Def II: Mittlere (\pm SEM) Veränderung des EDSS $-0,11 \pm 0,060$ vs. $0,04 \pm 0,079$; $p=0,074$ und des MSFC $-0,01 \pm 0,030$ vs. $-0,09 \pm 0,041$; $p=0,010$.

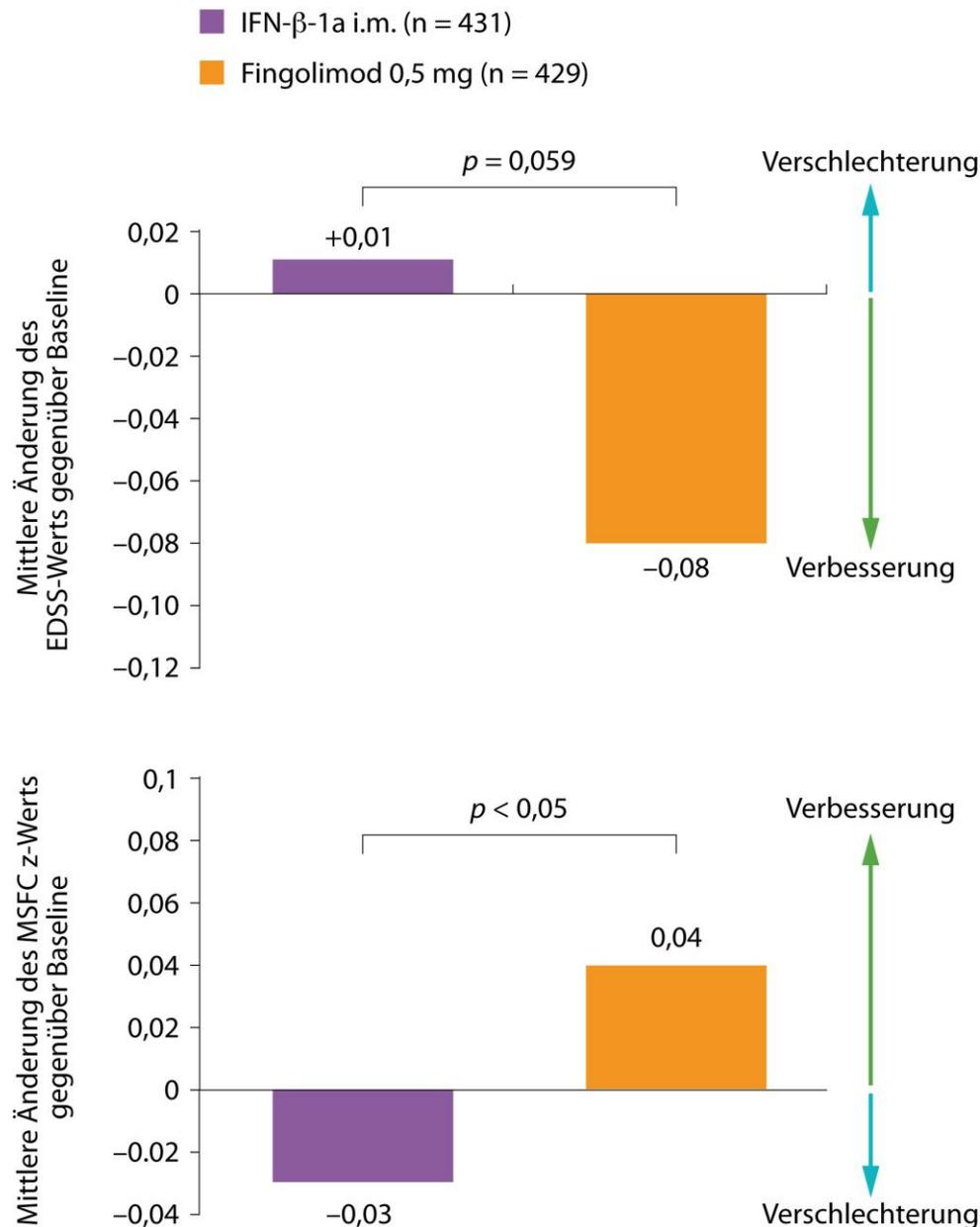


Abbildung 4-15: Änderung gegenüber Studienbeginn des EDSS-Werts und des MSFC z-Punktwerts in Monat 12 in TRANSFORMS (ITT-Population)

Bei den EDSS-Werten stellt eine negative Änderung eine Verbesserung dar und eine positive Änderung eine Verschlechterung. Bei den MSFC z-Punktwerten stellt eine negative Änderung eine Verschlechterung dar und eine positive Änderung eine Verbesserung. Die p -Werte für Fingolimod versus IFN-β-1a i. m. wurden mithilfe des Rangsummentests nach Wilcoxon berechnet.

Endpunkte zur gesundheitsbezogenen Lebensqualität

Die mittlere Änderung des aus dem EQ-5D ermittelten gesundheitsbezogenen Lebensqualitätsindex zeigt abgesehen von der Subpopulation der Patienten mit schnell fortschreitender MS keinen Unterschied zwischen Fingolimod und Placebo beziehungsweise

der zweckmäßigen Vergleichstherapie IFN- β -1a i. m. Für die Subpopulation der Patienten mit schnell fortschreitender MS war die mittlere Änderung des Lebensqualitätsindex signifikant unterschiedlich für Fingolimod im Vergleich zu Placebo, wobei sich Fingolimod-Patienten in ihrer gesundheitsbezogenen Lebensqualität verbesserten und Placebo-Patienten verschlechterten. Tendenziell ähnlich fiel auch der Vergleich Fingolimod zu IFN- β -1a i. m. aus, allerdings war der Unterschied hier nicht signifikant.

Die Bewertung der gesundheitsbezogenen Lebensqualität durch den Patienten mittels der visuellen Analogskala (VAS) des EQ-5D zeigte einen nahezu konsistenten Trend zugunsten von Fingolimod. Die Änderung gegenüber Baseline zeigte dabei zumeist eine geringere Verschlechterung oder sogar ausgeprägtere Verbesserung des Zustandes unter Fingolimod im Vergleich zu Placebo oder IFN- β -1a i. m. Für die Population der IFN Non-Responder war diese Verbesserung statistisch signifikant für Fingolimod im Vergleich zu IFN- β -1a i. m.. Fingolimod scheint hinsichtlich der Beeinflussung der gesundheitsbezogenen Lebensqualität von Vorteil gegenüber der zweckmäßigen Vergleichstherapie IFN- β -1a i. m. zu sein, mit einem besonders deutlichen Vorteil für die IFN Non-Responder.

Diese Ergebnisse sind aber aufgrund der geringen Spezifität des EQ-5D für RRMS (indikationsübergreifendes Instrument) eher von untergeordneter Aussagekraft. Signifikante Unterschiede sind mit diesem Instrument in dieser Indikation kaum zu erwarten. Zudem konnte keine Aussage getroffen werden, inwiefern eine orale Therapie im Gegensatz zu einer Injektionstherapie hinsichtlich der Lebensqualität von Vorteil ist. Das Double-Dummy Design der TRANSFORMS-Studie lässt das nicht zu, da jeder Patient sowohl eine Kapsel als auch eine Injektion erhalten hat.

Unerwünschte Ereignisse

Dank umfangreicher Studien im Bereich RRMS ist das Sicherheitsprofil von Fingolimod gut beschrieben. In FREEDOMS und TRANSFORMS wies die Dosis Fingolimod 0,5 mg gegenüber der Dosis Fingolimod 1,25 mg ein besseres Nutzen/Risiko-Profil auf. Aus diesem Grund wurde die 0,5 mg Dosierung zur Zulassung eingereicht. Im Folgenden werden nur Sicherheitsdaten aus der 0,5 mg Dosisgruppe sowie der Vergleichsgruppen berichtet. Sicherheitsdaten zur höheren Dosierung sind den Studienberichten zu entnehmen. In den relevanten Subpopulationen wurden keine bedeutsam von der Gesamtpopulation abweichenden Daten festgestellt. Die Konfidenzintervalle aus den Auswertungen der Subpopulationen sind relativ weit. Ein Vergleich der Daten zwischen Gesamt- und relevanten Subpopulationen ist daher schwer interpretierbar.

Fingolimod wird von Patienten mit schubförmig verlaufender Multipler Sklerose gut vertragen und die Studienabbruchrate ist gering. Das Sicherheitsprofil der zur Zulassung beantragten Dosis von 0,5 mg ist durch entsprechende therapiebegleitende Maßnahmen gut kontrollierbar. Für die 0,5 mg-Dosierung lag die Nebenwirkungsrate über 2 Jahre auf Placebo-Niveau (FREEDOMS: Behandlungsabbruch aufgrund eines UE bei 7,5 % im

Fingolimod 0,5 mg-Arm versus 7,7 % im Placebo-Arm). Die meisten UE waren von milder bis moderater Ausprägung, schwerwiegende UE (SUE) traten nur selten auf (FREEDOMS: 10,1 % im Fingolimod 0,5 mg Arm versus 13,4 % im Placebo-Arm). Aus dem indirekten Vergleich mit Glatirameracetat geht hervor, dass die Rate an Behandlungsabbrüchen aufgrund von UE unter Glatirameracetat um das 1,7 fache im Gegensatz zu Fingolimod erhöht ist. Fingolimod bietet tendenziell einen Sicherheitsvorteil gegenüber Glatirameracetat.

Die in der Fachinformation von Fingolimod empfohlenen therapiebegleitenden Maßnahmen dienen der Vermeidung potenzieller Nebenwirkungen und der optimalen Versorgung von MS-Patienten. Größtenteils werden die notwendigen Untersuchungen im Rahmen der Versorgungsroutine durchgeführt.

Häufige UE

Häufige UE in beiden Studien (Sicherheitspopulation) waren Kopfschmerzen (RR und 95 %-KI in FREEDOMS 1,08 (0,85; 1,38) und TRANSFORMS 1,14 (0,89; 1,47)) und Nasopharyngitis (RR und 95 %-KI in FREEDOMS 0,98 (0,79; 1,23) und TRANSFORMS 1,00 (0,77; 1,31)), die mit einer ähnlichen Häufigkeit von rund 20-28 % (Fingolimod 0,5 mg, Placebo und IFN- β -1a i. m.) berichtet wurden. Infektionen der oberen Atemwege wurden bei 17,3 % der Patienten in der Studie FREEDOMS berichtet (Fingolimod 0,5 mg und Placebo, RR und 95 %-KI: 0,98 (0,73; 1,32)), jedoch betrug die berichtete Gesamtinzidenz in TRANSFORMS rund 7 % (Fingolimod 0,5 mg und IFN- β -1a i. m.; RR und 95 %-KI: 1,15 (0,70; 1,90)). Während der zweijährigen Studie FREEDOMS waren Kopfschmerzen, auffällige Leberfunktionstests, Grippe, Diarrhö, Husten und Rückenschmerzen diejenigen UE (Inzidenz > 10 %), die unter Fingolimod bei einem größeren Anteil Patienten berichtet wurden als unter Placebo (≥ 2 % Unterschied zwischen den Gruppen). Analog dazu wurden in der TRANSFORMS-Studie nur Kopfschmerzen (Inzidenz > 10 %) unter Fingolimod bei einem größeren Anteil Patienten berichtet als unter IFN- β -1a i. m. (≥ 2 % Unterschied zwischen den Gruppen), das 95 %-Konfidenzintervall des relativen Risikos zeigt keine Signifikanz (RR 1,14 (0,89; 1,47)). Auffallend war hier außerdem die hohe Rate von grippeähnlichen Symptomen in der IFN- β -1a i. m. Gruppe, die gegenüber 0,5 mg Fingolimod um das 10-fache erhöht war (36,9 % vs. 3,5 %). Das relative Risiko lag hier bei 0,09 für Fingolimod im Vergleich zu IFN- β -1a i. m. (95 %-KI: 0,06; 0,16). Dieser Unterschied zeigte sich auch in einer ca. 4-fachen Erhöhung der Pyrexie (17,9 % vs. 4,2 %; RR und 95 %-KI: 0,23 (0,14; 0,39)) sowie einer mehr als 3-fachen Erhöhung der Myalgien (10,2 % vs. 3,3 %; RR und 95 %-KI: 0,32 (0,18; 0,57)).

Bradyarrhythmie/AV-Überleitungsstörungen/ QT-Zeit-Verlängerungen:

In den Studien der Phase III wurde eine vorübergehende, in der Regel asymptotische Senkung der Herzfrequenz im Mittel um 8 Schläge pro Minute innerhalb einer Stunde nach der ersten Gabe von Fingolimod 0,5 mg beobachtet, die 4-5 Stunden nach Verabreichung das Maximum erreichte. Klinisch relevante Wirkungen auf die Herzfrequenz bildeten sich weitgehend innerhalb von 24 Stunden nach der ersten Gabe von Fingolimod zurück (3). Bei wiederholter Verabreichung wurden keine additiven Effekte festgestellt (4). Als SUE berichtete Bradykardien traten sowohl in FREEDOMS als auch in TRANSFORMS unter Fingolimod 0,5 mg häufiger auf als unter Vergleichstherapie (4 vs. 1 in FREEDOMS; 2 vs 0 in TRANSFORMS). Aber weder gegenüber Placebo noch gegenüber IFN- β -1a i. m. konnte ein signifikant erhöhtes Risiko für Bradykardien, AV-Überleitungsstörungen oder QT-Zeit-Verlängerungen abgeleitet werden.

Infektionen:

Angesichts der immunmodulatorischen Eigenschaften von Fingolimod könnte ein erhöhtes Risiko für Infektionen erwartet werden. Jedoch waren die Anteile der Patienten mit Infektionen insgesamt ähnlich für alle Behandlungsgruppen in den Studien TRANSFORMS (51%-52 %) und FREEDOMS (72 %), und die Infektionen waren meist von leichter oder mäßiger Intensität. Infektionen, für die in einer der Behandlungsgruppen aus dem 95%-Konfidenzintervall tatsächlich ein höheres Risiko abzuleiten ist sind lediglich Bronchitis 2,23 (1,23; 4,03) und Tinea versicolor 4,92 (1,08; 22,31), jeweils gegenüber Placebo. Gegenüber IFN- β -1a i. m. ist kein erhöhtes Risiko für eine bestimmte Infektionsart festzustellen. Es wurden keine opportunistischen Infektionen berichtet.

Blutdruck:

Geringe Anstiege des mittleren arteriellen Blutdrucks um 2-3 mmHg wurden während der ersten 6 Monate der Behandlung beobachtet. Diese Anstiege blieben für die Dauer der 12-monatigen Studie TRANSFORMS und der 24-monatigen Studie FREEDOMS stabil. Über die Dauer der Studien war aus dem 95 %-Konfidenzintervall weder gegenüber Placebo noch gegenüber IFN- β -1a i. m. ein erhöhtes Risiko für einen Anstieg des Blutdrucks unter Fingolimod ableitbar.

Leberwerterhöhungen:

Sowohl in TRANSFORMS als auch in FREEDOMS wurde eine Erhöhung der Leberwerte festgestellt. Es kam bei 8 % der Patienten unter Fingolimod 0,5 mg gegenüber 2 % der Placebo-Patienten zu einem Anstieg der Lebertransaminasen auf das 3-fache der Obergrenze

des Normalwerts (ULN) und darüber. Ein Anstieg auf mehr als das 5-fache der ULN trat bei 2 % der Fingolimod- und 1 % der Placebo-Patienten auf. In den Studien wurde Fingolimod abgesetzt, falls der Anstieg das 5-fache des ULN überstieg. Bei erneuter Exposition gegenüber Fingolimod traten bei einigen Patienten erneut erhöhte Lebertransaminasen-Werte auf, was auf einen Zusammenhang mit Fingolimod hindeutet. Die Mehrzahl dieser Erhöhungen trat innerhalb von 3-4 Monaten auf. Die Serumtransaminase-Werte normalisierten sich innerhalb von rund 2 Monaten nach Absetzen von Fingolimod. Das relative Risiko für Leberwerterhöhungen verschiedener Art ist unter Fingolimod gegenüber Placebo beziehungsweise IFN- β -1a i. m. um das 3- bis 4-fache erhöht.

Lymphopenie/ Leukopenie

Ein zentraler pharmakodynamischer Effekt von Fingolimod ist die dosisabhängige Reduktion der peripheren Lymphozytenzahl bis auf 20-30 % des Ausgangswertes. Dies ist auf das reversible Zurückhalten (Sequestrierung) von Lymphozyten im Lymphgewebe zurückzuführen. Aus der FREEDOMS-Studie ist unter Fingolimod ein erhöhtes Risiko für Lymphopenien und Leukopenien im Vergleich zu Placebo ableitbar. In der TRANSFORMS-Studie trat unter Fingolimod nur ein Fall einer Lymphopenie auf. Leukopenien traten unter Fingolimod häufiger auf, ein erhöhtes relatives Risiko gegenüber IFN- β -1a i. m. lässt sich aber nicht belegen.

Makulaödem:

Makulaödeme traten im gesamten Studienprogramm bei weniger als 1 % der mit Fingolimod behandelten Patienten auf. In FREEDOMS wurde unter Fingolimod 0,5 mg kein Makulaödem berichtet, in TRANSFORMS bei 0,5 % der Patienten. Alle berichteten Fälle von Makulaödem waren nach Absetzen von Fingolimod zumeist ohne die Notwendigkeit weiterer Interventionen reversibel.

Maligne Erkrankungen:

Insgesamt wurden maligne Erkrankungen in den beiden Phase III-Studien selten berichtet; die Fälle traten in TRANSFORMS häufiger unter Fingolimod (12 Patienten; 2,8 %) als unter IFN- β -1a i. m. auf (3 Patienten; 0,7 %), waren jedoch in FREEDOMS unter Fingolimod weniger häufig (5 Patienten; 1,2 %) als unter Placebo (10 Patienten; 2,4 %). Integrierte Sicherheitsdaten aus FREEDOMS und TRANSFORMS bis zu 12 Monaten zeigen, dass bösartige Neubildungen in ähnlichen Anteilen bei den mit Fingolimod behandelten Patienten (0,5 mg, 1,2 %) und den Patienten, die Vergleichsbehandlungen erhielten, auftraten (IFN- β -1a i. m. 0,5 %; Placebo 1,4 %). Ein erhöhtes Risiko wurde von den Zulassungsbehörden nicht gesehen.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Tabelle 4-1: Bewertung des Zusatznutzens innerhalb der relevanten Patientenpopulationen

	Gesamtpopulation gegen IFN- β -1a i. m.	Anwendungsgebiet a2 gegen IFN- β -1a i. m. (dargestellt über Population IFN Non-Responder Def I und Def II)	Anwendungsgebiet b gegen IFN- β -1a i. m. (Schnell fortschreitende schwere RRMS)	Anwendungsgebiet a1 gegen Glatirameracetat (über indirekten Vergleich)
Reduziert die jährliche Schubrate	✓✓	✓✓	✓	✓✓
Reduziert das Schubrisiko bzw erhöht den Anteil schubfreier Patienten	✓✓	✓✓	✓	=
Reduziert die Anzahl der schweren Schübe	✓ ^a	✓ ^a	✓ ^a	Nicht bestimmbar
reduziert die Anzahl der hospitalisierungspflichtigen Schübe	✓	✓	✓	Nicht bestimmbar
Reduziert die Anzahl steroidpflichtiger Schübe	✓	✓	✓	Nicht bestimmbar
Reduziert das Risiko der Behinderungsprogression	✓	✓	Nicht bestimmbar	=
Verbessert die Behinderungsparameter EDSS MSFC	✓✓ ✓	✓ ✓✓	✓ ✓✓	Nicht bestimmbar
Verbessert die gesundheitsbezogene Lebensqualität	✓	✓	✓	Nicht bestimmbar
Bietet ein gutes Sicherheitsprofil	≈ (und Vermeidung grippeartiger Symptome)	≈ (und Vermeidung grippeartiger Symptome)	≈ (und Vermeidung grippeartiger Symptome)	✓
✓✓: signifikant überlegen ✓: tendenziell überlegen =: kein Unterschied nachweisbar ≈: vergleichbar -: unterlegen a: nur overall p-Wert ermittelt über milde, mittelschwere und schwere Schübe				

Für Fingolimod konnte ein medizinischer Nutzen und Zusatznutzen gegenüber den zweckmäßigen Vergleichstherapien gezeigt werden. Die relevanten Patientenpopulationen wurden soweit möglich abgebildet, eine Unterscheidung der IFN Non-Responder in therapietreue und nicht-therapietreue Patienten war mit der verfügbaren Datenlage nicht möglich (s. Abschnitt 4.2.5.1). Die ausgewerteten Patientenpopulationen enthalten jeweils beide Gruppen und sind daher für beide aussagekräftig. Der Zusatznutzen von Fingolimod gegenüber den zweckmäßigen Vergleichstherapien hinsichtlich patientenrelevanter Endpunkte wie der Reduktion der Schubrate und der Behinderungsprogression sowie das gut untersuchte Sicherheitsprofil rechtfertigen Fingolimod zur Therapie der RRMS in den relevanten Patientenpopulationen. Das Sicherheits- und Verträglichkeitsprofil von Fingolimod ist gut charakterisiert und stellt insgesamt auch eine für den Patienten subjektiv wahrnehmbare Verbesserung gegenüber der zweckmäßigen Vergleichstherapie dar. Der Zusatznutzen hinsichtlich der Wirksamkeit und dem Ausbleiben von alltagseinschränkenden applikationsbedingten Nebenwirkungen kann daher die Therapiemotivation und die damit einhergehende Therapieadhärenz entscheidend verbessern. Die Ergebnisse zum medizinischen Zusatznutzen erlauben die Schlussfolgerung, dass Fingolimod eine langfristige Freiheit von schwerwiegenden Symptomen möglich macht. Fingolimod zeigt eine nachhaltige und bisher nicht erreichte Verbesserung des therapielevanten Nutzens in allen relevanten Patientenpopulationen.

Der patientenrelevante Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie kann daher in allen relevanten Patientenpopulationen als erheblich beschrieben werden und kann aufgrund der Studie der Klasse 1-Evidenz, die der Bewertung zugrunde liegt, als gesichert angesehen werden.

Durch die Zulassung von Fingolimod steht den MS-Patienten eine weitere Therapieoption zur Verfügung, die viele der derzeitigen ungelösten Herausforderungen in der Therapie der MS adressiert. Der medizinische Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie hinsichtlich Schubratenreduktion und Behinderungsprogression sowie das gut untersuchte Sicherheitsprofil befürworten den Einsatz von Fingolimod zur Therapie der schubförmig remittierenden MS in den gemäß Zulassung in Frage kommenden Patientengruppen .

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

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.

Es soll nachgewiesen werden, dass Fingolimod 0,5 mg einmal täglich oral bei Patienten mit hochaktiver RRMS (d. h. IFN Non-Responder und Patienten mit schnell fortschreitender schwerer RRMS) hinsichtlich patientenrelevanter Endpunkte, also z. B. der schubbezogenen Endpunkte, der Endpunkte zur Behinderungsprogression und zur gesundheitsbezogenen Lebensqualität sowie den unerwünschten Ereignissen einen medizinischen Nutzen (Vergleich zu Placebo) beziehungsweise Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie (Vergleich zu IFN- β -1a i. m. 30 μ g einmal wöchentlich intramuskulär, beziehungsweise Glatirameracetat 20 mg einmal täglich subkutan) besitzt. Es soll dabei insbesondere geprüft werden, ob Patienten, die aufgrund mangelnder Therapietreue (Compliance) nicht auf IFN- β respondieren, durch Fingolimod einen Zusatznutzen gegenüber IFN- β -1a i. m. haben. Bei Patienten, die trotz ausreichender Therapietreue nicht auf IFN- β respondieren, soll ein Zusatznutzen gegenüber Glatirameracetat nachgewiesen werden. Der Zusatznutzen bei Patienten, die an einer schnell fortschreitenden schweren RRMS leiden, soll gegenüber IFN- β -1a i. m. nachgewiesen werden. Die Patientenpopulationen und die jeweiligen Vergleichstherapien sind durch ein Beratungsgespräch (1) zur zweckmäßigen Vergleichstherapie wie folgt definiert:

- Im Anwendungsgebiet zu a:

- 1) Für Patienten, die nicht auf einen vollständigen und angemessenen normalerweise mindestens ein Jahr andauernden Zyklus einer Beta-Interferon-Therapie angesprochen haben Glatirameracetat.
 - 2) Für Patienten, die noch keine ausreichende Therapie mit Beta-Interferonen erhalten haben Beta-Interferon 1a oder 1b.
- Im Anwendungsgebiet zu b: Beta-Interferon 1a oder 1b

Die Unterteilung des Anwendungsgebiets in die Untergruppen a und b ergibt sich dabei aus der zugelassenen Indikation. In der Zulassung wird zwischen Gruppen von Patienten wie folgt unterschieden:

Gilenya ist als krankheitsmodifizierende Monotherapie von hochaktiver schubförmig-remittierend verlaufender Multipler Sklerose bei folgenden Gruppen erwachsener Patienten angezeigt:

- Patienten mit hoher Krankheitsaktivität trotz Behandlung mit einem Beta-Interferon. Dabei kann es sich um Patienten handeln, die nicht auf einen vollständigen und angemessenen (normalerweise mindestens ein Jahr andauernden) Zyklus einer Beta-Interferon-Therapie angesprochen haben. Diese Patienten sollten während der Therapie im vorangegangenen Jahr mindestens einen Schub gehabt haben und sie sollten mindestens neun T2-hyperintense Läsionen im kranialen MRT oder mindestens eine Gadolinium anreichernde Läsion aufweisen. Ein Patient, der nicht auf die Therapie anspricht („Non-Responder“), lässt sich ebenso als ein Patient mit einer im Vergleich zum Vorjahr unveränderten oder vermehrten Schubrate oder anhaltend schweren Schüben definieren

oder

- Patienten mit rasch fortschreitender, schwerer schubförmig-remittierend verlaufender Multipler Sklerose, definiert durch zwei oder mehr Schübe mit Behinderungsprogression in einem Jahr, und mit einer oder mehr Gadolinium anreichernden Läsionen im MRT des Gehirns oder mit einer signifikanten Erhöhung der T2-Läsionen im Vergleich zu einer kürzlich durchgeführten MRT.

Unter a1 sind somit therapietreue Patienten, die auf IFN- β nicht respondieren (Bezeichnung: Non-Responder im Anwendungsgebiet a1, compliant) zu verstehen, unter a2 nicht-therapietreue Patienten, die auf IFN- β nicht respondieren (Bezeichnung: Non-Responder im Anwendungsgebiet a2, non-compliant). Unter b sind Patienten mit rasch fortschreitender schwerer RRMS zu verstehen (Bezeichnung: Patienten mit rasch fortschreitender schwerer RRMS im Anwendungsgebiet b).

Die Nachweise sollen soweit verfügbar mittels randomisierter kontrollierter Studien (Evidenzstufe 1b) erfolgen. Sofern keine direkten Vergleichsstudien verfügbar sind, kann sich der Nachweis auf indirekte Vergleiche auf Basis von RCT stützen.

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.

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.

Tabelle 4-2: Ein- und Ausschlusskriterien für Studien zur Bewertung des Nutzens und Zusatznutzens

Parameter	Einschlusskriterien	Ausschlusskriterien
Indikation (direkter Vergleich)	RRMS: mindestens eine der in der Zulassung von Fingolimod genannten Patientengruppen muss Teil der Studienpopulation sein	Ausschließlich andere Indikationen als RRMS ausschließlich Patienten mit anderen Verlaufsformen der RRMS, als die in der Zulassung von Fingolimod genannten (z. B.: milde Verlaufsformen, seltene Schübe)
Indikation (indirekter Vergleich)	RRMS	Ausschließlich andere Indikationen als RRMS ausschließlich Patienten mit anderen Verlaufsformen der RRMS, als die in der Zulassung von Fingolimod genannten (z. B.: milde Verlaufsformen, seltene Schübe)
Patientenpopulation	Erwachsene Patienten ab 18 Jahren	Pädiatrische Patienten, gesunde Probanden
Intervention (direkte)	Fingolimod 0,5 mg einmal täglich	Studien mit ausschließlich

Parameter	Einschlusskriterien	Ausschlusskriterien
Vergleichsstudien)	oral	anderen Fingolimod-Dosierungen außer 0,5 mg
Vergleichstherapie (direkte Vergleichsstudien mit Fingolimod)	Placebo IFN- β -1a i. m. 30 μ g einmal wöchentlich (medizinischer Zusatznutzen) Glatirameracetat 20 mg einmal täglich s. c. (es sollen nur die gemäß Fachinformation empfohlenen Standarddosierungen einbezogen werden)	Bei aktiv-kontrollierten Studien ausschließlich andere Vergleichspräparate als IFN- β -1a i. m. oder Glatirameracetat (z. B. IFN- β -1b, Natalizumab, Azathioprin, Mitoxantron, Cyclophosphamid) Ausschließlich andere Dosierungen als die in den Einschlusskriterien genannten Vergleichstherapie darf nicht identisch mit der Intervention sein, z.B. bei Studien zur Fragestellung Frühtherapie vs. Verzögerte Therapie
Intervention (Studien für Netzwerkvergleich, indirekter Vergleich)	IFN- β -1a i. m. 30 μ g einmal wöchentlich IFN- β -1a s. c. 44 μ g dreimal wöchentlich IFN- β -1b s. c. 250 μ g jeden zweiten Tag Glatirameracetat 20 mg einmal täglich subkutan (es sollen nur die gemäß Fachinformation empfohlenen Standarddosierungen einbezogen werden) (5-9)	Andere Präparate als die unter den Einschlusskriterien genannten (z. B. Natalizumab, Azathioprin, Mitoxantron, Cyclophosphamid) Ausschließlich andere Dosierungen als die in den Einschlusskriterien genannten
Vergleichstherapie (Studien für Netzwerkvergleich, indirekter Vergleich)	Placebo IFN- β -1a i. m. 30 μ g einmal wöchentlich IFN- β -1a s. c. 44 μ g dreimal wöchentlich IFN- β -1b s. c. 250 μ g jeden zweiten	Bei aktiv-kontrollierten Studien ausschließlich andere Vergleichspräparate als die unter den Einschlusskriterien genannten (z. B. Natalizumab, Azathioprin, Mitoxantron, Cyclophosphamid) Andere Dosierungen als die in

Parameter	Einschlusskriterien	Ausschlusskriterien
	<p>Tag</p> <p>Glatirameracetat 20 mg einmal täglich s. c.</p> <p>(es sollen nur die gemäß Fachinformation empfohlenen Standarddosierungen einbezogen werden) (5-9)</p>	den Einschlusskriterien genannten
Endpunkte	<p>Morbidität:</p> <p>Schubbezogene Endpunkte (z. B. jährliche Schubrate, Anzahl Schübe, Schubcharakteristika, Zeit bis zum ersten Schub, Anzahl Patienten mit Schüben)</p> <p>Behinderungsprogression (z. B. Zeit bis zur bestätigten Behinderungsprogression, Anteil Patienten mit Behinderungsprogression, EDSS-/MSFC-Änderungen über die Zeit)</p> <p>Unerwünschte Ereignisse (z. B. Anzahl UE, SUE, Behandlungsabbrüche aufgrund UE)</p> <p>Gesundheitsbezogene Lebensqualität</p> <p>Mortalität</p>	Studien mit ausschließlich Surrogat-Parametern (z. B. MRT-bezogene Parameter)
Studientyp	Randomisierte kontrollierte Studien	<p>Extensionsstudien, Dosis-Reduktionsstudien, nicht-interventionelle Studien, andere Datenerhebungen</p> <p>Meta-Analysen</p>
Berichtstyp	Studienbericht vorliegend	<p>Nur Abstract vorliegend</p> <p>Review-Artikel, Surveys, Notes,</p>

Parameter	Einschlusskriterien	Ausschlusskriterien
	Publikation vorliegend	Letter, Editorial, Errata
Studiendauer	Studien mit Dauer \geq 12 Monate	Studien mit Dauer $<$ 12 Monate

In die Bewertung des medizinischen Nutzens und Zusatznutzens werden randomisierte kontrollierte Studien (RCT: Evidenzstufe 1b) eingeschlossen. Sofern RCT vorliegen, sind Studien niedrigeren Evidenzgrades für die Bewertung nicht weiter relevant. Ein Studienbericht oder eine Volltext-Publikation muss vorliegen. Review-Artikel und Meta-Analysen sollen nicht einbezogen werden. Extensionsstudien und Dosisreduktionsstudien sollen nicht einbezogen werden, da diese meist nicht auf eine Wirksamkeitsaussage ausgelegt sind, sondern der Weiterbehandlung der Patienten dienen.

Die für die Nutzen- und Zusatznutzenbewertung heranzuziehenden Studien müssen hinsichtlich Indikation und Patientenpopulation für die gemäß Zulassung in Frage kommenden Patienten aussagekräftig sein. Es sind daher nur Studien relevant, die RRMS-Patienten einschließen. Die Patientengruppen, die in der zugelassenen Indikation genannt werden, also IFN Non-Responder und Patienten mit schnell fortschreitender RRMS, dürfen nicht aus den Studien ausgeschlossen worden sein. Die Ein- und Ausschlusskriterien in den einzelnen Studien müssen also mindestens einen Teil der in der Zulassung von Fingolimod genannten Patientengruppen abdecken, da sonst keine Aussagekraft für die Therapie mit Fingolimod gegeben ist. Gemäß dieser Kriterien konnten keine Studien für einen indirekten Vergleich von Glatirameracetat mit Fingolimod identifiziert werden, die die IFN Non-Responder definieren. Entweder war eine Vortherapie von vornherein ausgeschlossen, oder die Vortherapien konnten keine IFN beinhalten, weil die Studien vor Einführung der IFN-Präparate rekrutiert hatten. Ein indirekter Vergleich mit Bezug auf die für Fingolimod relevanten Populationen ist daher nicht durchführbar. Die Suche wurde für diese Fragestellung daher auf RRMS-Patienten erweitert. Die Studienpopulationen in den RCT sollen entsprechend der Zulassung für Fingolimod aus Erwachsenen (ab 18 Jahre) bestehen.

Die Studien, die für die Bewertung relevant sind, müssen mindestens einen patientenrelevanten Endpunkt enthalten (Morbidität, Mortalität und gesundheitsbezogene Lebensqualität). Studien, die ausschließlich Surrogatparameter als Endpunkte aufführen, sind nicht für die Bewertung relevant.

Nur Studien mit Fingolimod in der zugelassenen Dosierung von 0,5 mg sind relevant für den direkten Vergleich. Vergleichstherapien, die für die Studiauswahl relevant sind, sind Placebo um den medizinischen Nutzen von Fingolimod zu zeigen und IFN- β -1a i. m. sowie Glatirameracetat, um den medizinischen Zusatznutzen zu zeigen (siehe Begründung zur zweckmäßigen Vergleichstherapie in Modul 3). Da die Recherche keine direkten Vergleichsstudien mit Glatirameracetat ergeben hat, wurde ein indirekter Vergleich notwendig. Für diese Netzwerkanalyse wurden Studien mit IFN- β -Präparaten und

Glatirameracetat gewählt - mit den jeweils zugelassenen Dosierungen und Applikationsschemata in der empfohlenen Standarddosierung und jeweils im Vergleich untereinander oder gegen Placebo. Die Standarddosierung wird gewählt, da für die Fingolimod-Therapie nur eine Dosierung zugelassen ist, ein Vergleich mit anderen Dosierungen der Komparatoren könnte das Ergebnis stark verzerren.

Es wurden ausschließlich Studien mit einer Studiendauer ≥ 12 Monate in die Analyse eingeschlossen. Aufgrund der Tatsache, dass es sich um eine chronische Erkrankung handelt, wird eine kürzere Studiendauer als nur bedingt aussagekräftig gesehen. Eine Aussage zur Behinderungsprogression oder zur Entwicklung der Schubrate lässt sich in kürzeren Studien nicht mit ausreichender Sicherheit zur Beantwortung der obigen Fragestellung treffen. Insbesondere für den Endpunkt der jährlichen Schubrate ist eine kürzere Studiendauer nicht valide, da dann vorhandene Daten auf ein Jahr extrapoliert werden müssen, was zu einer unangemessenen Unschärfe der Aussage im Hinblick auf die im Dossier erörterte Fragestellung führt.

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 in folgenden Fällen erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt:

- Für das zu bewertende Arzneimittel wurde eine Zulassung in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wurde, mindestens 12 Monate vor dem Zeitpunkt der

Dossiereinreichung erteilt. Dabei sind auch Zulassungen außerhalb Deutschlands bzw. Europas relevant.

- Es werden indirekte Vergleiche zur Bewertung des medizinischen Zusatznutzens herangezogen.

Sofern eine bibliografische Literaturrecherche erforderlich ist, soll diese mindestens in den Datenbanken MEDLINE und EMBASE sowie in den Cochrane-Datenbanken 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. Eine gemeinsame Suche nach Studien zu mehreren Fragestellungen (z. B. direkt vergleichende Studien sowie Studien für einen indirekten Vergleich) ist möglich. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

Beschreiben Sie nachfolgend, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Geben Sie dabei an, ob Sie eine gemeinsame Suche für mehrere Fragestellungen durchgeführt haben (z. B. gemeinsame Suche nach Studien mit dem zu bewertenden Arzneimittel und Studien zur zweckmäßigen Vergleichstherapie für indirekte Vergleiche). Begründen Sie Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, ob bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.

Die Durchführung einer bibliographischen Literaturrecherche war erforderlich für den direkten Vergleich (erster Zulassungstermin international: 17. August 2010 in Russland, siehe Modul 2) und den indirekten Vergleich (notwendig, da keine direkten Vergleichsstudien gegen Glatirameracetat für Fingolimod gefunden wurden).

Die Suche erfolgte am 25. Juli 2011. Es wurde getrennt gesucht nach: Fingolimod-Studien, IFN- β -Studien gegen Placebo, IFN- β -Studien im aktiven Vergleich untereinander, Glatirameracetat-Studien gegen Placebo und Glatirameracetat-Studien im Vergleich zu IFN- β -Präparaten. Durchsuchte Datenbanken waren: EMBASE, Medline, Cochrane Database of Systematic Reviews, Cochrane Central Register of Clinical Trials. Die Publikationssprache wurde nicht eingeschränkt. Eine zeitliche Einschränkung wurde nicht vorgenommen. Die zeitlichen Angaben beziehen sich auf den maximal verfügbaren Zeitraum (Medline ab 1948, EMBASE ab 1980). Die Cochrane Database auf Systematic Reviews wurde ab 2005 durchsucht, das entspricht aufgrund der regelmäßigen Aktualisierung der vorhandenen Reviews einer vollständigen Suche. Das Cochrane Register of Clinical Trials wurde von Quartal 3 2011 an rückwärts komplett durchsucht.

Zur Beurteilung der Daten wurde an einigen Stellen des Moduls 4 auf bibliographische Quellen verwiesen, die mittels Handsuche und aus internen Datenbanken identifiziert wurden.

4.2.3.3 Suche in Studienregistern

Eine Suche in öffentlich zugänglichen Studienregistern ist grundsätzlich durchzuführen, um sicherzustellen, dass laufende Studien sowie abgeschlossene Studien von Dritten vollständig identifiziert werden.

Die Suche soll mindestens in den Studienregistern clinicaltrials.gov, clinicalstudyresults.org sowie über das International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden.

Die Suche soll in jedem Studienregister einzeln und mit einer für das jeweilige Studienregister adaptierten Suchstrategie durchgeführt werden. Die Suche soll abgeschlossene, abgebrochene und laufende Studien erfassen. Eine gemeinsame Suche nach Studien zu mehreren Fragestellungen (z. B. direkt vergleichende Studien sowie Studien für einen indirekten Vergleich) ist möglich. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

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

In den öffentlichen Suchregistern clinicaltrials.gov, clinicalstudyresults.org und über das International Clinical Trials Registry Platform Search Portal wurde nach klinischen Studien mit Fingolimod beziehungsweise FTY720 sowie in der Indikation Multiple Sklerose gesucht. Die durchgeführten Suchstrategien entsprechen den obigen Vorgaben und sind in Anhang 4-B dokumentiert. Bei der Recherche wurden keine generellen Einschränkungen vorgenommen.

4.2.3.4 Selektion relevanter Studien

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

Die durch die Suche in bibliographischen Datenbanken und der Studienregister-Suche identifizierten Zitate wurden anhand ihres Titels und, sofern vorhanden, ihres Abstracts von zwei Reviewern unabhängig voneinander hinsichtlich ihrer Relevanz bewertet (anhand der in Abschnitt 4.2.2 aufgeführten Kriterien). Publikationen, die beide Reviewer als potenziell relevant erachteten, wurden anhand ihres Volltextes auf Relevanz geprüft. Zitate, die nur ein Reviewer als potenziell relevant einstufte, wurden von beiden Reviewern nochmals gesichtet und anschließend nach Diskussion entweder als irrelevant bezeichnet oder ebenfalls anhand ihres Volltextes auf Relevanz geprüft.

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-G):

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 des Behandlers
- 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 des Verzerrungspotenzials eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Das Verzerrungspotenzial der selektierten RCT wurde folgendermaßen beurteilt: Zunächst wurde für die Bewertung der Studien das Verzerrungspotenzial anhand der endpunktübergreifenden Aspekte als „niedrig“ oder „hoch“ eingestuft. Weiterhin wurden die endpunktspezifischen Aspekte beurteilt. Details zum Verzerrungspotenzial sind den Bewertungsbögen in Anhang 4-G 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)¹. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-² bzw. STROBE-Statements³ folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

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

Die Beschreibung des Designs und der Methodik der RCT erfolgte anhand der durch das CONSORT-Statement vorgegebenen Standards. Die Informationen zu den Items 2b bis 14 wurden jeweils dem Studienprotokoll und dessen Amendments bzw. den relevanten Publikationen entnommen. Die ausführliche Darstellung pro Item findet sich pro Studie in Anhang 4-F.

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 Endpunkte, zu denen Ergebnisse dargestellt werden. Begründen Sie, wenn Sie von den oben benannten Vorgaben abgewichen sind.

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

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

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

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). Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.

Patientencharakteristika

Die im Folgenden beschriebenen Studienergebnisse wurden an erwachsenen Patienten mit der Diagnose RRMS erhoben. Dieses Studienkollektiv ist als geeignet für die Beurteilung des medizinischen Nutzens und Zusatznutzens von Fingolimod zu sehen, da es die Patientengruppe, die für die Behandlung mit Fingolimod in Frage kommt, umfasst.

Zur Beurteilung der gemäß Zulassung relevanten Patientengruppen wurden post-hoc Subpopulationen definiert und separat ausgewertet. Definiert sind im Zulassungstext Patienten, die nicht auf IFN- β respondieren und Patienten mit schnell fortschreitender schwerer RRMS. IFN- β Non-Responder sind dabei Patienten, die trotz IFN- β -Therapie im vorangegangenen Jahr mindestens einen Schub hatten und mindestens neun T2-hyperintense Läsionen oder mindestens 1 Gd+ Läsion aufweisen. Eine alternative Definition ergibt sich über: eine im Vergleich zum Vorjahr unveränderte oder vermehrte Schubrate oder anhaltend schwere Schübe. Patienten mit rasch fortschreitender schwerer RRMS sind definiert über zwei oder mehr Schübe mit Behinderungsprogression in einem Jahr und mit einer oder mehr Gd+ Läsion oder einer signifikanten Erhöhung der T2-Läsionen im Vergleich zu einem kürzlich durchgeführten MRT. Aus dem Beratungsgespräch zur zweckmäßigen Vergleichstherapie (s. Modul 3) ergibt sich bei den IFN- β Non-Respondern eine weitere Unterscheidung nach der Therapietreue (Compliance) der Patienten. Die Auswertung der Subpopulationen aus den verfügbaren RCT folgte soweit möglich diesen Kriterien. Die Unterscheidung in therapietreue und nicht-therapietreue Patienten war nicht möglich, da die Baseline-Daten der Studien diese Unterscheidung nicht enthielten. Die Subpopulationen der IFN Non-Responder enthalten somit beide Gruppen (compliant und non-compliant). Die Bewertung muss daher auf dieser best-verfügbaren Evidenz basieren. Gemäß der beiden genannten Möglichkeiten der Definition eines Non-Responders wurden zwei Subpopulationen gebildet, IFN Non-Responder Def I (Definition über Schubrate und MRT) und IFN Non-Responder Def II (Definition über den Vergleich zweier Jahre) und separat dargestellt. Nicht berücksichtigt werden konnte bei der Gruppe B das Kriterium der „anhaltend schweren Schübe“, da diese Information aus den Studiendaten nicht verfügbar war. Die Subpopulation der Patienten mit schnell fortschreitender schwerer RRMS wurde ohne Berücksichtigung der Kriterien der Assoziation einer Behinderungsprogression und der Erhöhung der Anzahl T2-Läsionen gebildet, da diese Daten nicht verfügbar waren. Die gebildeten Subpopulationen sind die best-verfügbare Evidenz zur Beantwortung der Fragestellung. Studiendaten der TRANSFORMS und der FREEDOMS-Studie werden jeweils für die ITT und die drei gebildeten Subpopulationen dargestellt.

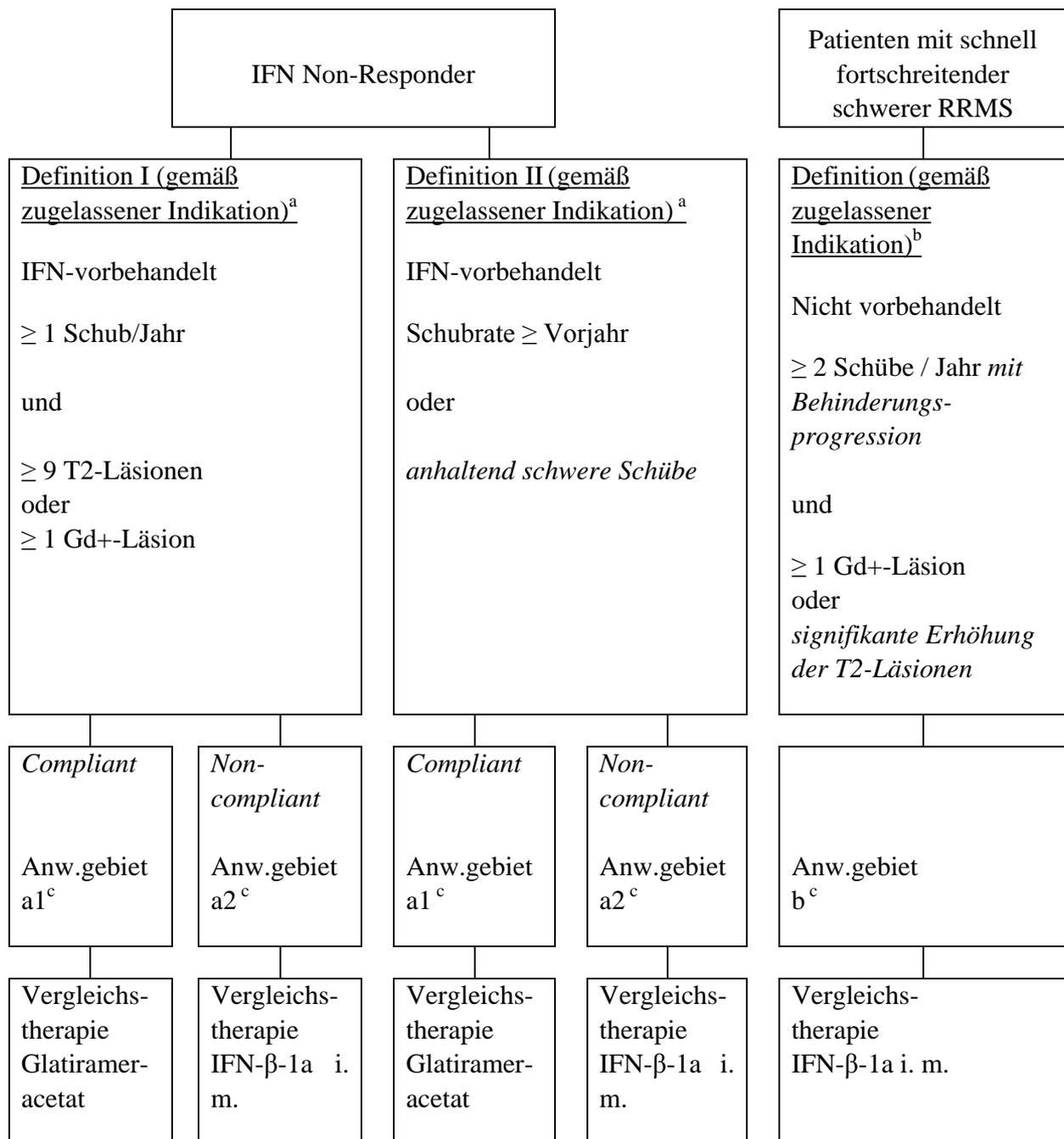
Für den indirekten Vergleich konnten keine Subpopulationen gebildet werden, da die in den dafür herangezogenen Publikationen Informationen dafür nicht ausreichten. Gemäß der vordefinierten Kriterien konnten keine Studien für einen indirekten Vergleich von

Glatirameracetat mit Fingolimod identifiziert werden, die die IFN Non-Responder definieren. Entweder war eine Vortherapie von vornherein ausgeschlossen, oder die Vortherapien konnten keine IFN beinhalten, weil die Studien vor Einführung der IFN-Präparate rekrutiert hatten. Ein indirekter Vergleich mit Bezug auf die für Fingolimod relevanten Populationen ist daher nicht durchführbar. Die Suche wurde für diese Fragestellung daher auf alle erwachsenen RRMS-Patienten erweitert. Die verwendeten Publikationen sind zur Fragestellung die best-verfügbare Evidenz.

Im Ergebnisteil werden die ITT-Population und die Subpopulationen anhand folgender Charakteristika beschrieben (ausgewählt aufgrund der Relevanz hinsichtlich der Definition der Subpopulationen und Subgruppen):

- Alter
- Geschlecht
- Schubrate zu Baseline
- Krankheitsschwere gemäß EDSS
- Patienten ohne Gd+-Läsionen zu Baseline
- Dauer der Erkrankung
- Patienten ohne Vorbehandlung

Übersicht der Subpopulationen, die in Modul 4 dargestellt werden im Vergleich zu den in der Zulassung und dem Beratungsgespräch zur zweckmäßigen Vergleichstherapie festgelegten Kriterien:



a: entspricht den im Modul 4 dargestellten Subpopulationen IFN Non-Responder Def I und II und umfasst jeweils die Anwendungsgebiete a1 und a2; Definition nur teilweise erfüllt (kursiv: nicht aus Rohdaten auszulesen)

b: entspricht der im Modul 4 dargestellten Subpopulation Patienten mit schnell fortschreitender schwerer RRMS; Definition nur teilweise erfüllt (kursiv: nicht aus Rohdaten auszulesen)

c: gemäß Beratungsgespräch mit dem G-BA am 01. Juli 2011 (1)

Endpunkte

Die durchgeführten klinischen Studien wurden unter Beachtung der CHMP-Guideline CPMP/EWP/561/98 Rev 1 („Guideline on Clinical Investigation of Medicinal Products for the Treatment of Multiple Sclerosis“) sowie der Direktive 75/318/EEC und den ICH Topics, E1, E4, E9, E10 konzipiert und durchgeführt (10-15).

Die CHMP-Guideline CPMP/EWP/561/98 Rev 1 (10) sieht für die Untersuchung der Wirksamkeit in konfirmatorischen Studien neuer krankheitsmodifizierender Therapeutika zur Behandlung der schubförmigen Multiplen Sklerose als primäre Wirksamkeitsparameter die Veränderung der Behinderungsprogression oder die Schubrate vor. Als Methode der Wahl zur Untersuchung der Behinderungsprogression sollte der EDSS herangezogen werden. Therapieversagen oder Progression sollte vordefiniert werden, z. B. als das Erreichen eines spezifischen Grades der Behinderung oder als anhaltende (nach 3 oder 6 Monaten bestätigte) Verschlechterung relevanten Ausmaßes (Zunahme um 1 Punkt bei EDSS $\leq 5,5$; Zunahme um 0,5 Punkte bei EDSS $> 5,5$). Weitere Bewertungen anhand neurologischer Skalen oder quantitativer Neuro-Performance-Tests (z. B. MSFC) sollten in Ergänzung durchgeführt werden. Zur Erfassung der Veränderung in der Schubhäufigkeit ist die jährliche Schubrate ein akzeptierter Parameter, wobei eine präzise prospektive Definition von Schub im Allgemeinen (Auftreten, Dauer, Abklingen, Schwere etc.) notwendig ist. Als sekundäre Wirksamkeitsendpunkte sind je nach Wahl des primären Endpunktes wiederum Veränderung der Behinderungsprogression oder die Schubrate, sowie MRT-Parameter und weitere Parameter, wie neurologische oder kognitive Skalen, vorgesehen.

Gemäß IQWiG Methodenpapier {IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2008 556 /id} ist patientenrelevant, was das Gefühl, die Wahrnehmung von Funktionen und Aktivitäten oder das Überleben des Patienten betrifft. Im Speziellen sind demgemäß patientenrelevante Zielgrößen die Mortalität, Morbidität und die gesundheitsbezogene Lebensqualität. Der Aufwand, den die Behandlung oder die Erkrankung an sich für den Patienten mit sich bringt oder die Behandlungszufriedenheit können berücksichtigt werden.

Zur Bewertung des medizinischen Nutzens und Zusatznutzens von Fingolimod 0,5 mg werden die Ergebnisse der folgenden Endpunkte (soweit verfügbar) dargestellt:

- Jährliche Schubrate
- Anteil Patienten ohne Schübe
- Schubcharakteristika (Anzahl milder, mittelschwerer, schwerer Schübe, Anzahl hospitalisierungspflichtiger Schübe, Anzahl steroidpflichtiger Schübe)
- Anteil Patienten ohne Behinderungsprogression
- Mittlere Änderung des EDSS
- Mittlere Änderung des MSFC

- Mittlere Änderung des EQ-5D
- Unerwünschte Ereignisse (Anzahl UE, SUE, Behandlungsabbrüche aufgrund UE, relevante UE)

Zur Beurteilung des medizinischen Nutzens und Zusatznutzens wurden somit ausschließlich Endpunkte gewählt, die nach Einschätzungen des CHMP für das Studienziel geeignet sind. Somit können alle im Folgenden dargestellten Endpunkte als valide angesehen werden. Alle Endpunkte sind zudem patientenrelevant, wie im Folgenden dargestellt wird:

Wie in Modul 3 beschrieben, geht ein Schub mit einer Verschlechterung der klinischen Symptomatik einher. Für die Dauer des Schubes ist der Patient je nach Symptomatik in verschiedenen Bereichen stark eingeschränkt und gegebenenfalls sind zusätzliche therapeutische Maßnahmen erforderlich. Zudem ergibt sich eine Progression der Behinderung unter anderem aus der unvollständigen Reversibilität der Schubsymptomatik. Die Reduktion der klinischen Schubhäufigkeit ist ein entscheidendes Therapieziel. Die schubbezogenen Endpunkte „Jährliche Schubrate“, „Anteil Patienten ohne Schübe“ und die „Schubcharakteristika“ sind patientenrelevant, weil sie den Gesundheitszustand und damit die Morbidität bewerten.

Der Grad der Behinderung spiegelt direkt das Ausmaß der Einschränkungen, denen ein Patient ausgesetzt ist, wider. Über diese Angaben wird der Endpunkt „Behinderungsprogression“ ermittelt. Eine Steigerung des Behinderungsgrades wird, wie oben beschrieben, als Progression gewertet und zeigt damit eine Einschränkung des Patienten an. Der Endpunkt „Behinderungsprogression“ und die Endpunkte zur Änderung des EDSS und des MSFC sind patientenrelevant, weil sie den Gesundheitszustand und damit die Morbidität bewerten. Die Bewertungsskalen EDSS und MSFC sind anerkannt. Die EDSS basiert auf der DSS (Disability Status Scale), die 1955 von John Kurtzke entwickelt und validiert wurde (17). Die DSS misst Verschlechterung in Schritten von einem Punkt. Kurtzke verfeinerte die DSS indem er die Skala in Schritten von halben Punkten aufteilte (18). Eine Validierungsstudie zum EDSS fehlt, bis heute ist der EDSS aber weltweit der Goldstandard zur Messung der Behinderungsprogression in der MS und wird in der europäischen Leitlinie zu Wirksamkeitsstudien in der MS als Instrument empfohlen (10). Der EDSS erfasst weder kognitive Funktionen noch die Funktion der oberen Extremitäten. In der CHMP-Leitlinie (10) wird daher die zusätzliche Verwendung eines weiteren Instruments, z. B. des MSFC empfohlen. Der Fragebogen MSFC ist ein kombiniertes Messinstrument bestehend aus drei Komponenten für die objektive und quantitative Erfassung der neurologischen Funktion beziehungsweise Funktionseinschränkung der oberen und unteren Extremität sowie des kognitiven Funktionssystems. Der MSFC wurde Ende der 90er Jahre von Cutter et al. entwickelt und validiert (19).

Die gesundheitsbezogene Lebensqualität ist vom IQWiG {IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2008 556 /id} als patientenrelevanter Endpunkt definiert. Sie wird in Form von Fragebögen erfasst anhand derer der Patient den Gesundheitszustand und die Auswirkungen auf die gesundheitsbezogene Lebensqualität bewertet. Der EQ-5D ist ein Werkzeug zur Erfassung des Gesundheitszustandes und der

Auswirkungen auf die gesundheitsbezogene Lebensqualität, der Endpunkt zur Änderung des EQ-5D ist damit patientenrelevant. Allerdings ist der EQ-5D ein unspezifisches Instrument ohne Indikationsbezug. Der EQ-5D wurde von Fisk et al. für MS validiert (20). Eine Validierung des EQ-5D mit einem „Time Trade off“ für die deutsche Bevölkerung wurde von Greiner et al. publiziert (21). Die EQ-5D-Index-Daten, die hier im Dossier vorgelegt werden, basieren auf einem Werte-Set für das Vereinigte Königreich, da dies für die Studienauswertung definiert war.

Unerwünschte Ereignisse haben direkten Einfluss auf den Gesundheitszustand des Patienten und sind damit als Endpunkt patientenrelevant. Als relevante UE wurden Bradykardie/AV-Überleitungsstörungen, Infektionen, Hypertension, Leberwerterhöhungen, Makulaödem, Malignitäten und QT-Zeit-Verlängerung definiert. Ausgewählt wurden diese Ereignisse, weil sie entweder in der Fachinformation von Fingolimod genannt werden oder speziell für Immunmodulatoren relevant sind.

In den Studien TRANSFORMS und FREEDOMS wurden neben den oben aufgeführten Scores auch der uFIS (undimensional fatigue impact scale) und der PRIMUS (patient-reported indices in multiple sclerosis) berichtet. Beide Werkzeuge wurden für diese Studien entwickelt und sind wertvolle Werkzeuge, um die Auswirkungen der Krankheit und der Behandlung auf das tägliche Leben und die gesundheitsbezogene Lebensqualität zu bewerten. Als Standard sind sie allerdings noch nicht etabliert, daher wird auf die Darstellung verzichtet.

4.2.5.3 Meta-Analysen

Sofern die vorliegenden Studien dazu geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst werden. Dabei sollen verschiedene Studientypen nicht in einer Meta-Analyse zusammengefasst werden. Für Meta-Analysen soll die im Folgenden beschriebene Methodik eingesetzt werden.

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten⁴ erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär anhand des Odds Ratio durchgeführt werden. In begründeten

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

Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet⁵ werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität⁶ erfolgen. Ist die Heterogenität der Studienergebnisse nicht bedeutsam, soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

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

Bei den beiden vorliegenden Phase III-Studien handelt es sich um verschiedene Studientypen (aktiv- beziehungsweise Placebo-kontrolliert), daher erscheint eine Zusammenfassung der Studien TRANSFORMS und FREEDOMS mittels Meta-Analyse nicht sinnvoll.

4.2.5.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse sollen Sensitivitätsanalysen hinsichtlich methodischer Faktoren durchgeführt werden, wenn sich die Studien bezüglich dieser Faktoren unterscheiden. 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.

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

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

Die beiden Zulassungsstudien FREEDOMS und TRANSFORMS waren gemeinsam im Rahmen des Zulassungsprogramms konzipiert worden und wiesen hinsichtlich methodischer Faktoren – abgesehen von der Wahl der Vergleichstherapie – praktisch keine Unterschiede auf.

Innerhalb der Studien FREEDOMS und TRANSFORMS wurden folgende Sensitivitätsanalysen durchgeführt:

Die aggregierte ARR und nach 3-Monaten bestätigte Behinderungsprogression war innerhalb beider Studien für die ITT-Population konsistent mit der PP-Population. Die ARR auf Patientenlevel war innerhalb beider Studien sowohl für die ITT-Population als auch für die PP-Population konsistent mit der aggregierten ARR (22;23).

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 sollten, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

- Geschlecht
- Alter
- Krankheitsschwere bzw. -stadium

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

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

Benennen Sie die durchgeführten Subgruppenanalysen. Begründen Sie die durchgeführten Subgruppenanalysen bzw. die Untersuchung von Effektmodifikatoren oder den Verzicht auf solche Analysen. Beschreiben Sie die für diese Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen. Begründen Sie die Wahl von Trennpunkten, wenn quantitative Merkmale kategorisiert werden. Verwenden Sie dabei nach Möglichkeit die in dem jeweiligen Gebiet gebräuchlichen Einteilungen und begründen Sie etwaige Abweichungen. Kennzeichnen Sie in einzelnen Studien a priori geplante Subgruppenanalysen.

In den Studienprotokollen der beiden zur Bewertung herangezogenen RCT waren a priori keine Subgruppenanalysen geplant. Da die Subgruppenanalysen post-hoc durchgeführt wurden, sollte man zurückhaltend Aussagen über die Ergebnisse treffen, da sie möglicherweise nicht robust genug sind. Es ist nötig, die biologische Plausibilität der Ergebnisse, die Anzahl der durchgeführten Subgruppenanalysen, ob sie vordefiniert waren oder nicht, und die Fallzahl in Erwägung zu ziehen, bevor Rückschlüsse aus den Ergebnissen einer Subgruppenanalyse gezogen werden.

Im Ergebnisteil werden die Subgruppenanalysen zu Krankheitsschwere und Geschlecht dargestellt. Die Krankheitsschwere wird dabei dargestellt über den EDSS-Wert zu Studienbeginn. Als Trennpunkt wird ein Wert von 3,5 festgelegt ($\leq 3,5$ und $> 3,5$). Die Definition der Subgruppen erfolgte post-hoc. Der Trennpunkt gründet darauf, dass oberhalb eines EDSS von 3,5 (also ab Werten 4,0 oder höher) eine Einschränkung der Gehfähigkeit vorliegt, Werte bis einschließlich 3,5 bedeuten noch keine Einschränkung der Gehfähigkeit (18). Dieser Trennpunkt unterscheidet sich von den im Studienbericht dargestellten Trennpunkten, die aber auch post-hoc definiert waren.

Von der Darstellung weiterer Subgruppenanalysen wird abgesehen, da sie als nicht relevant und nicht aussagekräftig angesehen werden. Insbesondere beim Alter lässt sich kein sinnvoller Trennpunkt für Altersgruppen festlegen.

Analysiert wurden alle Endpunkte, die auch für die ITT-Population sowie die Subpopulation analysiert wurden. Lediglich für den Endpunkt „Unerwünschte Ereignisse“ werden keine Subgruppenanalysen dargestellt, da hiervon kein zusätzlicher Erkenntnisgewinn erwartet werden kann. Eine Auswertung einzelner relevanter UE-Kategorien ist aufgrund der ohnehin schon geringen Ereigniszahlen nicht weiter aussagekräftig.

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Nicht adjustierte indirekte Vergleiche (d. h. Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator) stellen dabei keine valide Analysemethode dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich⁷. Komplexe Verfahren für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als „Mixed-Treatment-Comparison(MTC)-Meta-Analysen“⁸, „Multiple-Treatment-Meta-Analysen“⁹ oder auch „Netzwerk-Meta-Analysen“¹⁰, sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem

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

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

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

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

Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist¹¹.

Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen¹².

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

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

Indirekte Vergleiche werden zur Beurteilung des medizinischen Zusatznutzens für die Patientengruppe der therapietreuen IFN Non-Responder herangezogen (zweckmäßige Vergleichstherapie Glatirameracetat). Direkte Vergleichsstudien von Fingolimod gegen Glatirameracetat existieren nicht.

Es wurde daher eine Netzwerkanalyse der in Modul 3 genannten zweckmäßigen Vergleichstherapien (Basistherapeutika) durchgeführt.

Es konnte keine separate Auswertung für die Population der IFN Non-Responder durchgeführt werden. Die Publikationen boten dafür nicht die Datengrundlage. Gemäß der in Abschnitt 4.2.2 genannten Kriterien wurden keine Studien mit Glatirameracetat identifiziert, die die IFN Non-Responder definieren. Entweder war eine Vortherapie von vornherein ausgeschlossen, oder die Vortherapien konnten keine IFN beinhalten, weil die Studien vor Einführung der IFN-Präparate rekrutiert hatten. Ein indirekter Vergleich mit Bezug auf die für

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

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

Fingolimod relevanten Populationen ist daher nicht durchführbar. Die im Dossier dargestellte Netzwerkanalyse bezieht sich daher auf RRMS-Patienten allgemein und ist für die Fragestellung des Vergleichs von Fingolimod mit Glatirameracetat die best-verfügbare Evidenz.

Die im Weiteren beschriebene Methodik wurde dem Bericht über die Netzwerkanalyse entnommen (24).

Folgende Endpunkte wurden analysiert:

- Jährliche Schubrate (Mixed Treatment Comparison, Poisson)
- Anteil schubfreier Patienten (Mixed Treatment Comparison, binomial)
- Nach 3-Monaten bestätigte Behinderungsprogression (Mixed Treatment Comparison, binomial)
- Behandlungsabbrüche aufgrund von UE (Mixed Treatment Comparison, binomial)

Endpunkte mit einer binomialen Verteilung wurden folgendermaßen analysiert:

Für jeden binomialen Endpunkt wurde ein “mixed log-binomial” Modell aufgesetzt, i steht dabei für die jeweilige Studie und j für die Behandlung. Eine allgemeingültige Version des Modells lautet:

- $y_{ij} \sim \text{binomial}(n_{ij}, p_{ij})$
- $\log(p_{ij}) = t_j + s_i + st_{ij} + b \cdot x_{ij}$

wenn y_{ij} die Anzahl der Patienten ist, die den Endpunkt erreicht haben und n_{ij} die Gesamtzahl der Patienten in der Studie i mit der Behandlung j wird angenommen, dass

- s_i und st_{ij} zufällige Effekte sind, die einer Normalverteilung folgen (im Mittel “0” mit unbekannter Varianz)
- x_{ij} eine Kovariate ist (oder Kovariate-Matrix bei mehr als einer Kovariaten), die allen Patienten in der Studie i mit Behandlung j gemeinsam ist (z. B. mittleres Alter in der Behandlungsgruppe)
- t_j der Logarithmus der gesamten Ereigniswahrscheinlichkeit für die Behandlung j mit dem Mittel x_{ij} ist.

Das relative Risiko unter Behandlung a versus Behandlung b ist dann

$$p_a/p_b = \exp(t_a - t_b).$$

Der zufällige Effekt s_i berücksichtigt die Korrelation zwischen den Responsvariablen der Patienten innerhalb der jeweiligen Studie. Der zufällige Effekt st_{ij} berücksichtigt die

Korrelation der Responsvariablen beliebiger zwei Patienten innerhalb eines Behandlungsarms innerhalb der jeweiligen Studie.

Der Einschluss des zufälligen Effektes st_{ij} hängt von der Zusammensetzung der Daten ab. Wenn die Zahl der Beobachtungen kleiner als die Zahl der zu schätzenden Modellparameter ist, kann der Einschluss des zufälligen Effektes st_{ij} nicht aufrechterhalten werden und dieser zufällige Effekt wird deshalb im Modell nicht berücksichtigt. Dies ist das Szenario für diese die vorliegenden Meta-Analysen.

Folgender Programmcode (PROC GLIMMIX code) wurde für die binomiale “Mixed Treatment Comparison” verwendet (r = Anzahl Patienten, die den Endpunkt erreicht haben und n = Anzahl der Patienten in der Studie/Behandlungsarm).

```
proc glimmix data = outcome_data itdetails;
  class trt study;
  model r/n = trt xcovar / noint link = log solution;
  random intercept / subject = study solution;
  estimate "trt-a vs trt-b" trt 1 -1 0 0 ... / exp cl;
run;
```

Endpunkte, die einer Poisson-Verteilung folgen, wurden folgendermaßen analysiert:

Eine mixed Poisson-Regression wurde für jeden Poisson-verteilten Endpunkt aufgesetzt. i steht dabei für die jeweilige Studie und j für die Behandlung. Eine allgemeingültige Version des Modells lautet:

- $y_{ij} \sim \text{Poisson}(n_{ij}p_{ij})$
- $\log(p_{ij}) = t_i + s_j + st_{ij} + b \cdot x_{ij}$

wenn y_{ij} für die Anzahl der Ereignisse steht und n_{ij} die Expositionszeit der Patienten in der Studie i mit Behandlung j . wird angenommen, dass

- s_i und st_{ij} zufällige Effekte sind, die einer Normalverteilung folgen (im Mittel “0” mit unbekannter Varianz)
- x_{ij} eine Kovariate ist (oder Kovariate-Matrix bei mehr als einer Kovariaten), die allen Patienten in der Studie i mit Behandlung j gemeinsam ist (z. B. mittleres Alter in der Behandlungsgruppe)
- t_j der Logarithmus der Gesamtinzidenz für Behandlung j ist, wenn gilt $x_{ij} = 0$

Die relative Rate für Behandlung a versus Behandlung b ist dann $p_a/p_b = \exp(t_a - t_b)$.

Der zufällige Effekt s_i berücksichtigt die Korrelation zwischen den Responsvariablen der Patienten innerhalb der jeweiligen Studie. Der zufällige Effekt st_{ij} berücksichtigt die Korrelation der Responsvariablen beliebiger zwei Patienten innerhalb eines Behandlungsarms innerhalb der jeweiligen Studie.

Der Einschluss des zufälligen Effektes σ_{ij} hängt von der Zusammensetzung der Daten ab. Wenn die Zahl der Beobachtungen kleiner als die Zahl der zu schätzenden Modellparameter ist, kann der Einschluss des zufälligen Effekts σ_{ij} nicht aufrechterhalten werden und dieser zufällige Effekt wird deshalb im Modell nicht berücksichtigt. Dies ist das Szenario für diese die vorliegenden Meta-Analysen.

Folgender Programmcode (PROC GLIMMIX code) wurde verwendet (r = Ereignisse und $\logpy = \log$ (total person-time) für die jeweilige Studie/Behandlungsarm).

```
proc glimmix data = outcome_data itdetails;
  class trt study;
  model r = trt xcovar
  / noint distribution = poisson offset = logpy;
  random intercept / subject = study solution;
  estimate "trt-a vs trt-b" trt 1 -1 0 0 ... / exp cl;
run;
```

Bewertung der Eignung des Mixed Treatment Comparison Modells

Die Eignung des Modells kann informell über die Schätzer aus einer reduzierten Chi-Quadrat-Statistik bewertet werden (Chi-Quadrat dividiert durch die Freiheitsgrade im Modell). Schätzer nahe bei 1.0 stehen für ein gut geeignetes Modell, besonders im Hinblick auf die spezifizierten Zufallseffekte (s. Tabelle 4-3).

Tabelle 4-3: Bewertung des Modells der Netzwerkanalyse

Endpunkt	reduziertes Chi-Quadrat
Nach 3 Monaten bestätigte Behinderungsprogression	1.18
Jährliche Schubrate	1.35
Behandlungsabbrüche aufgrund von UE	1.38
Schubfreie Patienten	1.41

Quelle: RTI Bericht (24)

Zur Konsistenzprüfung erfolgt ein qualitativer Vergleich der Größe und Richtung der Effektschätzer für Glatirameracetat vs Placebo beziehungsweise Fingolimod vs. Placebo aus den Originalpublikationen/ Studienberichten mit den Effektschätzern aus den MTC-Modellen.

Sensitivitätsanalysen wurden nicht durchgeführt, da die Datengrundlage der Publikationen dafür nicht geeignet ist.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

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

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

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

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

Benennen Sie in der nachfolgenden Tabelle alle Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war.

Von den randomisierten kontrollierten Studien, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden, sollen die Studien, deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle randomisierten kontrollierten Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.

Benennen Sie nur randomisierte, kontrollierte Studien, 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“, Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

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

Studien- bezeichnung ^a	Zulassungs- studie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
FTY720D1201	ja (für Japan)	abgeschlossen	6 Monate	Fingolimod 0,5 mg Fingolimod 1,25 mg Placebo
FTY720D2201	nein	abgeschlossen	6 Monate	Fingolimod 1,25 mg Fingolimod 5,0 mg Placebo
FTY720D2301 FREEDOMS	ja	abgeschlossen	24 Monate	Fingolimod 0,5 mg Fingolimod 1,25 mg Placebo
FTY720D2302 TRANSFORMS	ja	abgeschlossen	12 Monate	Fingolimod 0,5 mg Fingolimod 1,25 mg IFN-β-1a i. m. 30 µg
FTY720D2306 INFORMS	nein	laufend	36+ Monate	Fingolimod 1,25 mg ^b Placebo
FTY720D2309 FREEDOMS II	nein	laufend	24 Monate	Fingolimod 0,5 mg Fingolimod 1,25 mg ^b Placebo
FTY720D2320	nein	abgeschlossen ^c	3 Monate	Fingolimod 0,5 mg Placebo
FTY720DUS01 EPOC	nein	laufend	6 Monate	Fingolimod 0,5 mg „Standard-of-Care“-DMT
CFTY720DIT01 GOLDEN	nein	laufend	18 Monate	Fingolimod 0,5 mg IFN-β-1b s.c.
CFTY720DIT02 EPOC	nein	laufend	6 Monate	Fingolimod 0,5 mg IFN-β-1a, IFN-β-1b oder Glatirameracetat
<p>a: Nicht gelistet sind Phase I-Studien</p> <p>b: Im November 2009 wurden alle Patienten auf Fingolimod 0,5 mg umgestellt</p> <p>c: noch keine Auswertung/Studienbericht vorliegend</p> <p>Anmerkung: Eine Übersicht zu allen mit Fingolimod in der RRMS durchgeführten und laufenden Studien ist in Modul 5 hinterlegt (Modul5, Studienliste_Fingolimod.pdf).</p>				

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

Die Informationen in Tabelle 4-4 wurden zuletzt am 26. Juli 2011 aktualisiert, d. h. die Tabelle enthält den Studienstatus zu diesem Datum.

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

Tabelle 4-5: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
FTY720D1201	Studiendauer zu kurz (6-Monats-Studie)
FTY720D2201	untersucht nicht die zugelassene Dosierung Studiendauer zu kurz (6-Monats-Studie)
FTY720D2306 INFORMS	laufende Studie, andere Indikation (PPMS)
FTY720D2309 FREEDOMS II	laufende Studie
FTY720D2320	noch keine Auswertung vorliegend
FTY720DUS01	laufende Studie
FTY720DIT01	laufende Studie
FTY720DIT02	laufende Studie

Stellen Sie die abgebrochenen und laufenden Studien, die nicht in die Bewertung eingehen, in Anhang 4-D und Anhang 4-E dar.

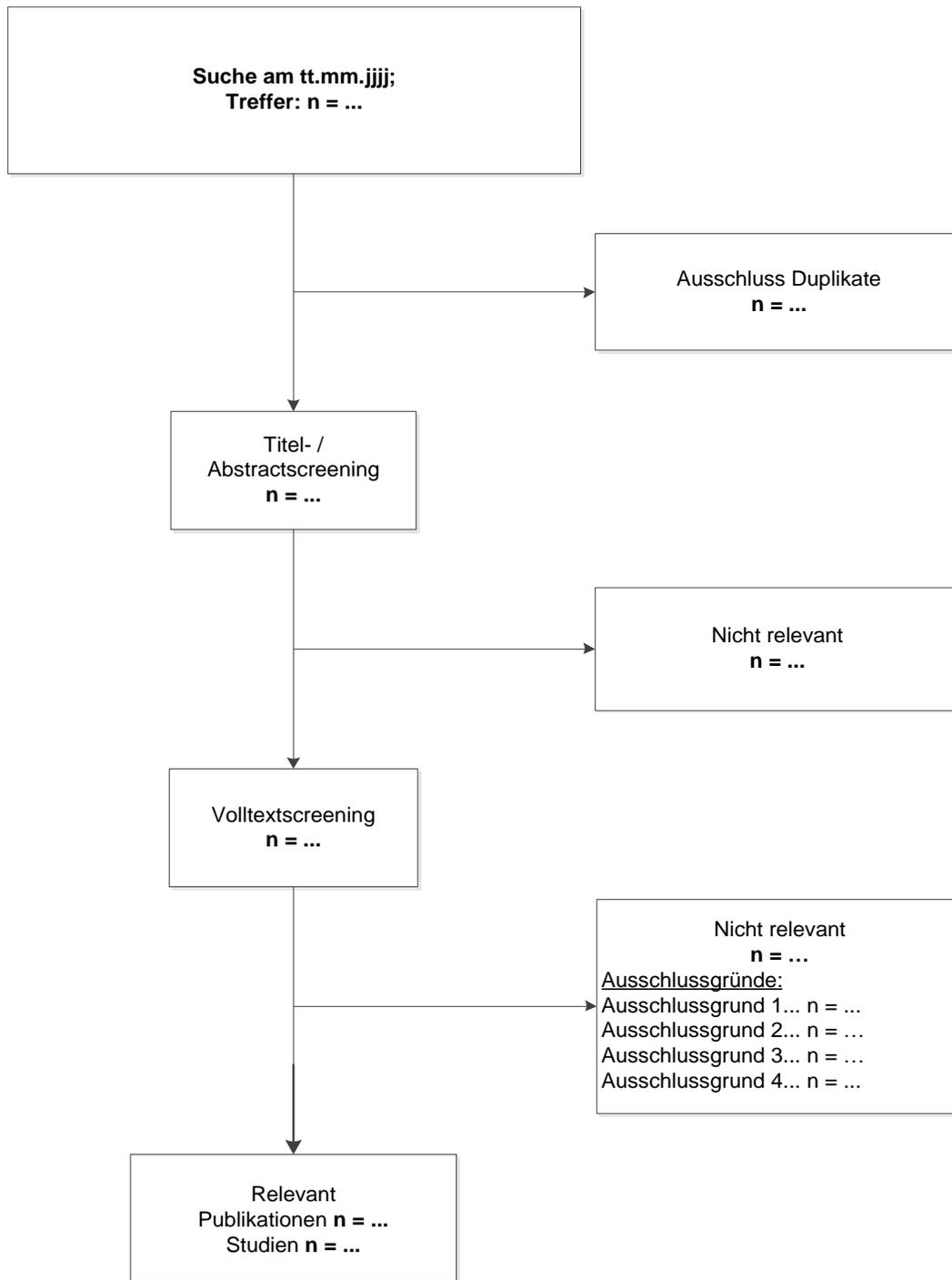
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

Sofern eine bibliografische Literaturrecherche durchgeführt wurde, beschreiben Sie nachfolgend das Ergebnis dieser Recherche. 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 Studien unter Nennung des Ausschlussgrunds in Anhang 4-C.

Falls die Recherche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel gemeinsam mit der Recherche zu einer anderen Fragestellung (z. B. zu indirekten Vergleichen) durchgeführt wurde, unterteilen Sie die Angaben zu relevanten Treffern und Studien entsprechend diesen Fragestellungen.

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

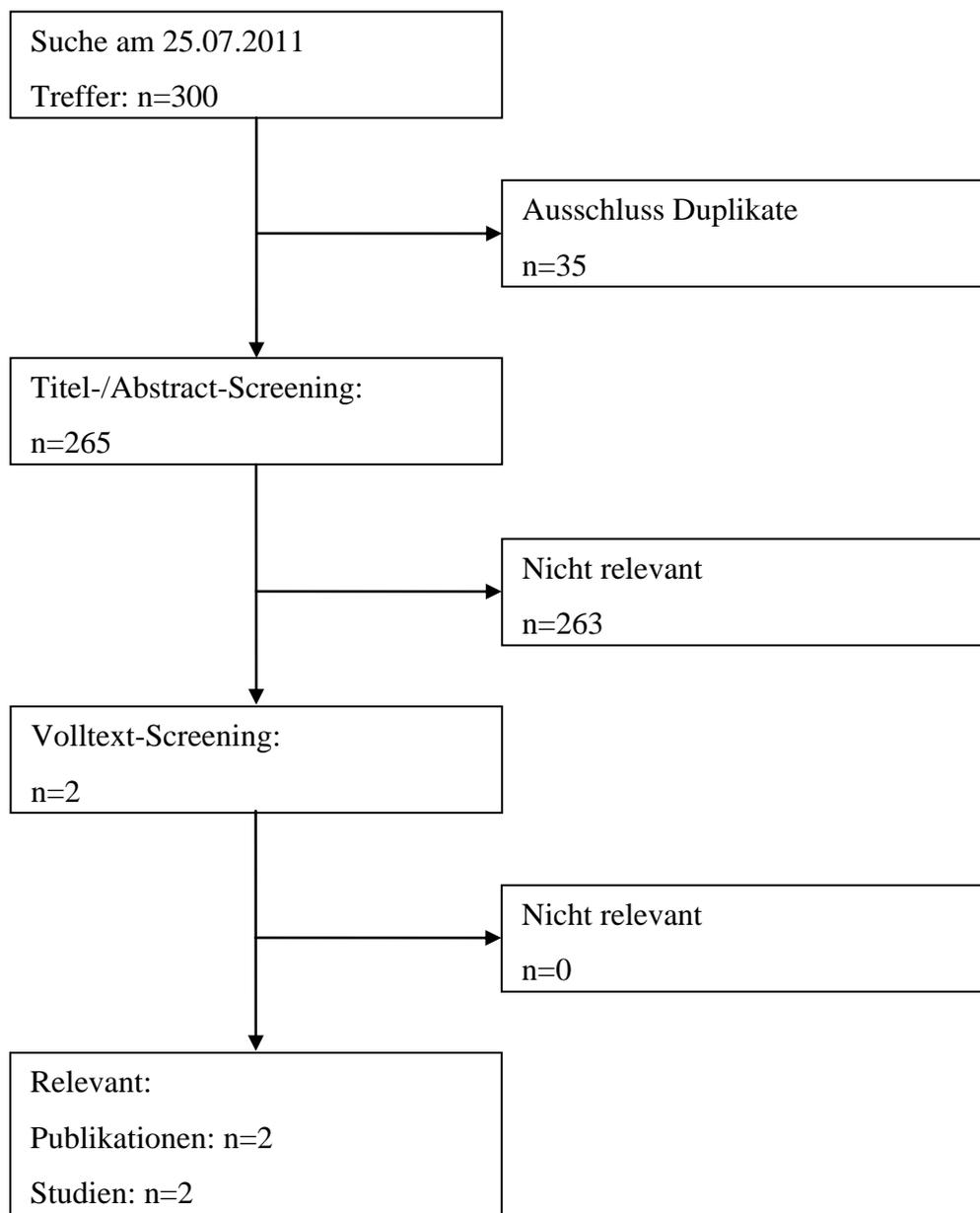
Orientieren Sie sich bei der Erstellung des Flussdiagramms an dem nachfolgenden Beispiel.



An dieser Stelle wird nur die Recherche zu Fingolimod-Studien beschrieben. Die Recherche zu IFN- β -Studien und Glatirameracetat-Studien wurde zum Zweck des indirekten Vergleichs durchgeführt und wird in Abschnitt 4.3.2.1.1.2 beschrieben.

Die Suchstrategie ist in Anhang 4-A dargestellt. Die Auswahl erfolgte anhand der in Abschnitt 4.2.2 beschriebenen Kriterien. Durchsuchte Datenbanken waren: EMBASE, Medline, Chochrane Database of Systematic Reviews, Cochrane Central Register of Clinical Trials. Die Publikationssprache wurde nicht eingeschränkt. Eine zeitliche Einschränkung wurde nicht vorgenommen. Die zeitlichen Angaben beziehen sich auf den maximal verfügbaren Zeitraum (Medline ab 1948, EMBASE ab 1980). Die Cochrane Database auf Systematic Reviews wurde ab 2005 durchsucht, das entspricht aufgrund der regelmäßigen Aktualisierung der vorhandenen Reviews einer vollständigen Suche. Das Cochrane Register of Clinical Trials wurde von Quartal 3 2011 an rückwärts komplett durchsucht.

Abbildung 4-16: Flussdiagramm der bibliographischen Suche zu Fingolimod



4.3.1.1.3 Studien aus der Suche in Studienregistern

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-4) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde (sofern eine solche durchgeführt 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 4-6: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Identifikationsorte (Name der Studienregister und Angabe der Zitate ^a)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d. ^b)
FTY720D2301 FREEDOMS	clinicaltrials.gov [NCT00289978], Registereintrag und Ergebnisse (25);, International Clinical Trials Registry Platform [NCT00289978], Registereintrag (26)	ja	ja
FTY720D2302 TRANSFORMS	clinicaltrials.gov [NCT00340834], Registereintrag und Ergebnisse (27) International Clinical Trials Registry Platform [NCT00340834], Registereintrag (28)	ja	ja
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. b: n. d. = bibliografische Literaturrecherche nicht durchgeführt.			

In Studienregistern wurden keine weiteren relevanten, in die Bewertung einzubeziehenden RCT identifiziert.

4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

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

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Datenquellen. 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 4-7: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
Placebo-kontrolliert						
FREED OMS	ja	ja	nein	ja [Novartis, Studienbericht und Zusatzauswertung] (22;29)	ja [NCT00289978] (25;26)	ja [Kappos et al., 2010](30)
aktiv-kontrolliert, zweckmäßige Vergleichstherapie(n)						
TRANS FORMS	ja	ja	nein	ja [Novartis, Studienbericht und Zusatzauswertung] (23;29)	ja [NCT00340834] (27;28)	ja [Cohen et al., 2010](31)
a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.						
b: Studie, für die der Unternehmer Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war.						
c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.						

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

4.3.1.2.1 Studiendesign und Studienpopulationen

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

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

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

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

Studie	Studien-design	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
FREEDOMS	RCT, doppelblind, parallel	Erwachsene mit RRMS ^a 1 Schub im letzten Jahr oder 2 Schübe in den letzten 2 Jahren EDSS 0 - 5,5	Fingolimod 1,25 mg (n=429); davon relevante Patienten ^b : 0 Fingolimod 0,5 mg (n=425); davon relevante Patienten ^b : n=112 (26,4 %) Placebo (n=418); davon relevante Patienten ^b : n=99 (23,7 %)	Screening: 1 Monat Behandlung: 24 Monate	Weltweit in 22 Ländern 1/2006 – 7/2009	Schübe; Behinderungsprogression, gesundheitsbezogene Lebensqualität, unerwünschte Ereignisse
TRANSFORMS	RCT, doppelblind, parallel	Erwachsene mit RRMS ^a 1 Schub im letzten Jahr, oder 2 Schübe in den letzten 2 Jahren EDSS 0 - 5,5	Fingolimod 1,25 mg (n=426); davon relevante Patienten ^b : 0 Fingolimod 0,5 mg (n=431); davon relevante Patienten ^b : n=207 (48,0 %) IFN-β-1a i. m. (n=435); davon relevante Patienten ^b : n=185 (42,5 %) und jeweiliges Placebo (Double-Dummy)	Screening: 1 Monat Behandlung: 12 Monate	Weltweit in 18 Ländern 5/2006 – 11/2008	Schübe; Behinderungsprogression, gesundheitsbezogene Lebensqualität, unerwünschte Ereignisse
<p>a: Die Studienpopulationen umfassen die gemäß Zulassung relevante Patientenpopulation. Im Folgenden wird neben der Darstellung der Charakteristika und Ergebnisse für die gesamte Studienpopulation auch auf die gemäß Zulassung relevanten Subpopulationen eingegangen. Zur Definition der Subpopulationen siehe Abschnitt 4.2.5.2.</p> <p>b: gemäß Zulassung relevante Patienten, Die Patientenzahl ergibt sich aus der Zahl der IFN Non-Responder zuzüglich der Patienten mit schnell fortschreitender schwerer RRMS. Dabei ergibt sich die Patientenzahl der IFN Non-Responder aus der Gesamtsumme der Patienten, die gemäß Def I oder Def II in Frage kommen.</p>						

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

Studie	Gruppe 1	Gruppe 2	(Gruppe 3)^a	Vorbehandlung der MS mit zugelassener Therapie n (%)
FREEDOMS	Fingolimod 0,5 mg 1x tgl.	Placebo 1xtgl.	Fingolimod 1,25 mg 1x tgl.	520 (40,9)
TRANSFORMS	Fingolimod 0,5 mg 1x tgl. + Placebo i. m. 1x/Woche	IFN- β -1a i. m. 30 μ g 1x/Woche + Placebo 1x tgl.	Fingolimod 1,25 mg 1x tgl. + Placebo i. m. 1x/Woche	752 (58,2)
a: Ergebnisse der Gruppe 3 werden nicht weiter dargestellt, da die untersuchte Dosierung nicht zur Zulassung eingereicht wurde				

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

Studie Gruppe	N	Alter (Jahre) (M ± SD)	Geschlecht w / m (%)	Dauer Er- krankung (Jahre) ^a (M ± SD)	Baseline EDSS (M ± SD)	Anzahl Schübe in den letzten 2 Jahren (M ± SD)	Teilnehmer ohne Gd-anreichernde T1 Läsionen n / N (%)	Teilnehmer ohne Vorbe- handlung ^c n (%)
FREEDOMS								
ITT	1272	37,1 ± 8,76	69,9 / 30,1	8,2 ± 6,60	2,40 ± 1,32	2,1 ± 1,19	782/1264 (61,9)	752 (59,1)
(Fingolimod 1,25 mg) ^b	429	37,4 ± 8,91	68,8 / 31,2	8,4 ± 6,86	2,41 ± 1,36	2,1 ± 1,25	257/424 (60,6)	259 (60,4)
Fingolimod 0,5 mg	425	36,6 ± 8,77	69,6 / 30,4	8,0 ± 6,60	2,30 ± 1,29	2,1 ± 1,13	263/424 (62,0)	244 (57,4)
Placebo	418	37,2 ± 8,60	71,3 / 28,7	8,1 ± 6,35	2,49 ± 1,29	2,2 ± 1,19	262/416 (63,0)	249 (59,6)
IFN Non- Resp.Def I	154	37,2 ± 9,34	69,5 / 30,5	9,4 ± 6,28	2,74 ± 1,38	2,4 ± 1,34	82 / 153 (53,6)	0
(Fingolimod 1,25 mg) ^b	45	37,7 ± 9,96	68,9 / 31,1	9,8 ± 7,49	2,80 ± 1,27	2,7 ± 1,53	23 / 44 (52,3)	0
Fingolimod 0,5 mg	57	37,6 ± 9,47	66,7 / 33,3	8,9 ± 5,93	2,43 ± 1,20	2,1 ± 1,01	32 / 57 (56,1)	0
Placebo	52	36,2 ± 8,73	73,1 / 26,9	9,5 ± 5,56	3,04 ± 1,59	2,5 ± 1,45	27 / 52 (51,9)	0
IFN Non- Resp.Def II	162	37,3 ± 9,16	67,3 / 32,7	9,1 ± 6,29	2,69 ± 1,38	2,2 ± 1,31	95 / 160 (59,4)	0
(Fingolimod 1,25 mg) ^b	48	37,4 ± 10,01	68,8 / 31,3	9,1 ± 7,37	2,60 ± 1,32	2,6 ± 1,66	26 / 46 (56,5)	0
Fingolimod 0,5 mg	60	37,7 ± 9,44	68,3 / 31,7	8,7 ± 5,75	2,44 ± 1,23	1,9 ± 0,92	37 / 60 (61,7)	0
Placebo	54	36,7 ± 8,14	64,8 / 35,2	9,6 ± 5,90	3,05 ± 1,53	2,2 ± 1,28	32 / 54 (59,3)	0
Pat. mit schnell fortschr. RRMS	127	32,7 ± 7,81	65,4 / 34,6	4,6 ± 5,11	2,11 ± 1,16	2,9 ± 1,18	127	127 (100)
(Fingolimod 1,25 mg) ^b	42	34,1 ± 8,10	61,9 / 38,1	5,0 ± 5,28	2,02 ± 1,21	3,0 ± 1,56	42	42 (100)
Fingolimod 0,5 mg	48	31,8 ± 7,38	68,8 / 31,3	3,5 ± 3,83	2,22 ± 1,21	2,9 ± 0,99	48	48 (100)

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

Placebo	37	32,4 ± 7,98	64,9 / 35,1	5,4 ± 6,18	2,08 ± 1,04	2,6 ± 0,86	37	37 (100)
TRANSFORMS								
ITT	1292	36,2 ± 8,50	67,3 / 32,7	7,4 ± 6,16	2,21 ± 1,30	2,2 ± 1,61	826/1264 (65,3)	540 (41,8)
(Fingolimod 1,25 mg) ^b	426	35,8 ± 8,39	68,8 / 31,2	7,3 ± 5,96	2,21 ± 1,31	2,2 ± 1,19	270/412 (65,5)	170 (39,9)
Fingolimod 0,5 mg	431	36,7 ± 8,81	65,4 / 34,6	7,5 ± 6,20	2,24 ± 1,33	2,3 ± 2,20	288/427 (67,4)	184 (42,7)
IFN-β-1a i. m.	435	36,0 ± 8,29	67,8 / 32,2	7,4 ± 6,33	2,19 ± 1,26	2,3 ± 1,22	268/425 (63,1)	186 (42,8)
IFN Non- Resp.Def I	453	36,9 ± 8,44	70,6 / 29,4	8,8 ± 6,00	2,44 ± 1,34	2,3 ± 2,15	304 / 452 (67,3)	453 (100)
(Fingolimod 1,25 mg) ^b	144	36,3 ± 8,18	74,3 / 25,7	8,6 ± 6,03	2,45 ± 1,44	2,2 ± 1,24	107 / 144 (74,3)	144 (100)
Fingolimod 0,5 mg	160	36,8 ± 8,95	72,5 / 27,5	8,7 ± 5,64	2,45 ± 1,35	2,5 ± 3,23	104 / 160 (65,0)	160 (100)
IFN-β-1a i. m.	149	37,7 ± 8,10	65,1 / 34,9	9,1 ± 6,36	2,41 ± 1,23	2,3 ± 1,15	93 / 148 (62,8)	149 (100)
IFN Non- Resp.Def II	456	37,0 ± 8,50	70,8 / 29,2	8,7 ± 6,02	2,42 ± 1,33	2,2 ± 2,12	320 / 453 (70,6)	456 (100)
(Fingolimod 1,25 mg) ^b	152	36,4 ± 8,40	71,1 / 28,9	8,3 ± 5,89	2,38 ± 1,42	2,1 ± 1,17	118 / 150 (78,7)	152 (100)
Fingolimod 0,5 mg	166	37,1 ± 8,84	73,5 / 26,5	8,5 ± 5,78	2,42 ± 1,33	2,4 ± 3,18	116 / 166 (69,9)	166 (100)
IFN-β-1a i. m.	138	37,6 ± 8,22	67,4 / 32,6	9,2 ± 6,44	2,46 ± 1,23	2,1 ± 1,09	86 / 137 (62,8)	138 (100)
Pat. mit schnell fortschr RRMS	85	32,8 ± 8,16	71,8 / 28,2	4,4 ± 5,29	1,92 ± 1,23	2,8 ± 0,97	85	85 (100)
(Fingolimod 1,25 mg) ^b	28	32,0 ± 8,66	64,3 / 35,7	5,5 ± 6,78	2,02 ± 1,22	3,1 ± 1,30	28	28 (100)
Fingolimod 0,5 mg	27	31,4 ± 8,52	74,1 / 25,9	3,7 ± 3,31	1,80 ± 0,96	2,6 ± 0,69	27	27 (100)
IFN-β-1a i. m.	30	34,8 ± 7,15	76,7 / 23,3	4,0 ± 5,16	1,95 ± 1,46	2,6 ± 0,76	30	30 (100)
a: Seit 1. Symptom								
b: Ergebnisse der Gruppe 3 werden nicht weiter dargestellt								
c: Patienten ohne Vorbehandlung: Patienten, die keinen der 5 zugelassenen Immunmodulatoren oder ein anderes MS-Medikament erhalten haben								

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

Die durchgeführten klinischen Studien wurden unter Beachtung der CHMP-Guideline CPMP/EWP/561/98 Rev 1 („Guideline for Medicinal Products for the Treatment of Multiple Sclerosis“) sowie der Direktive 75/318/EEC und den ICH Topics, E1, E4, E9, E10 konzipiert und durchgeführt (10-15).

Folgende Empfehlungen und Vorgaben sind Inhalt dieser Richtlinien:

Die Patientenauswahl sollte gemäß der CHMP-Guideline anhand der revidierten Kriterien nach McDonald erfolgen (32).

Die Studiendauer sollte bei explorativen parallelen doppelblinden Studien 6 Monate betragen, für konfirmatorische Studien werden mindestens 2 Jahre Studiendauer zum Wirksamkeitsnachweis als notwendig erachtet.

Ein doppelblinded Studiendesign sollte bevorzugt gewählt werden. In Fällen, in denen dies nicht möglich ist, sollte ein verblindeter Gutachter eingesetzt werden. Der Wirksamkeitsbeleg sollte auf der Überlegenheit in den Endpunkten Schubratenreduktion oder Verzögerung der Behinderungsprogression gegenüber Placebo oder jeglicher Vergleichstherapie beruhen.

Aktiv-kontrollierte, parallel-armige Studien, die das neue Arzneimittel mit einer bereits zugelassenen Therapiealternative vergleichen, sind notwendig, um einen Vergleich des Nutzen-Risiko-Verhältnisses zu ermöglichen. Oftmals ist es schwierig zu definieren, ob die gegenwärtige MS-Aktivität aus einem Therapieversagen der aktuellen Behandlung oder der Zunahme der Krankheitsaktivität unter Behandlung resultiert. Es wird daher empfohlen, dass in Studien einer Zweitlinientherapie ein Arm mit der bisherigen Therapie fortgeführt wird.

Um die Aufrechterhaltung der Wirksamkeit und weitere Informationen zur Langzeittherapie zu adressieren, sollte eine offene Extensionsstudie durchgeführt werden.

Die Sicherheit der neuen Therapie ist durch eine vollständige Dokumentation der unerwünschten Ereignisse während der klinischen Prüfung zu charakterisieren. Besonderes Augenmerk sollte auf die Untersuchung potenzieller unerwünschter Wirkungen, die für die untersuchte Wirkstoffklasse charakteristisch sind, wie z. B. Depression oder Krampfanfälle unter Interferonen oder schwere Infektionen oder Autoimmunerkrankungen unter der Behandlung mit Immunmodulatoren, gelegt werden. Organspezifische unerwünschte Ereignisse sollten sorgfältig überwacht werden (z. B. neurologische, kardiologische und hepatische Ereignisse).

Aufgrund der potenziellen Langzeitanwendung einer eingeführten MS-Behandlung sollte eine große und repräsentative Gruppe von Patienten für einen ausreichenden Zeitraum behandelt werden. Zum Zeitpunkt der Zulassung wird erwartet, Sicherheitsdaten über mindestens 2 Jahre einer aussagekräftigen Population aufzuweisen.

Design und Durchführung der im Nachfolgenden beschriebenen Studien entsprechen diesen Vorgaben. Bei einer der Zulassungsstudien von Fingolimod (**FREEDOMS**, Fingolimod Research Evaluating Effects of Daily Oral therapy in MS) handelt es sich um eine randomisierte, doppelblinde, Placebo-kontrollierte, multizentrische Studie über 24 Monate. In dieser Phase III-Studie wurden 1.272 Patienten mit RRMS einer 24-monatigen Behandlung mit einmal täglich Fingolimod 0,5 mg oder 1,25 mg Kapseln oder Placebo (im Verhältnis 1:1:1) zugewiesen. Ein unabhängiger Arzt überwachte die Studienteilnehmer über mindestens 6 Stunden nach Verabreichung der ersten oralen Dosis des Medikaments. Reguläre Studienvsiten erfolgten alle 3 Monate, innerhalb der ersten 3 Monate noch häufiger. Bei jedem Besuch im Rahmen der Studie wurden unerwünschte Ereignisse, Vitalzeichen und Routine-Laborwerte untersucht. Ein unabhängiges DSMB überwachte die Sicherheit und das Gesamt-Nutzen-Risiko-Verhältnis.

Im Rahmen der **TRANSFORMS Studie** (TRail Assessing injectable interferoN vS FTY720 Oral in Relapsing-remitting Multiple Sclerosis) wurde der medizinische Zusatznutzen von Fingolimod im Vergleich zu IFN- β -1a i. m. über 12 Monate untersucht. Die Auswahl von IFN- β -1a i. m. als Vertreter der Substanzklasse IFN- β für den direkten Vergleich mit Fingolimod in TRANSFORMS ist vor dem Hintergrund der als gleichwertig geltenden immunmodulatorisch wirksamen Basistherapien der RRMS bereits in Modul 3 begründet worden. In dieser multizentrischen, randomisierten, doppelblinden Double-Dummy-Parallelgruppenstudie der Phase III wurden 1.292 Patienten mit RRMS einer 12-monatigen Behandlung mit einmal täglich Fingolimod 0,5 mg oder 1,25 mg Kapseln oder IFN- β -1a i. m. 30 μ g wöchentlich (im Verhältnis 1:1:1) zugewiesen.

Das unabhängige DSMB untersuchte die Sicherheit und das Gesamt-Nutzen-Risiko-Verhältnis während der Studie. Ein unabhängiger, die erste Dosis verabreichender Arzt überwachte die Patienten in den 6 Stunden nach Verabreichung der ersten oralen Dosis der Prüfmedikation, um Vitalzeichen, EKG-Befunde und Bradykardie-Ereignisse zu bewerten. Reguläre Studienvsiten erfolgten alle 3 Monate, innerhalb der ersten 3 Monate noch häufiger. Bei jedem Besuch im Rahmen der Studie wurden unerwünschte Ereignisse, Vitalzeichen und Routine-Laborwerte untersucht.

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-11: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial auf Studienebene
			Patient	Behandler			
FREEDOMS	ja	ja	ja	ja	nein	nein	niedrig
TRANSFORMS	ja	ja	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial auf Studienebene sowohl für FREEDOMS als auch TRANSFORMS als „niedrig“ eingestuft, da jeweils eine adäquate Erzeugung der Randomisierungssequenz vorlag, die Gruppenzuteilung verdeckt erfolgte, sowohl Patient als auch Behandler verblindet waren (doppelblind), für die unterschiedlichen Interventionen in der TRANSFORMS ein Double-Dummy Design gewählt wurde und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte gab.

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 4-12: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Schübe		Behinderungsprogression			Lebensqualität		UE UE, SUE, Studienabbrüche, relevante UE
	Jährliche Schubrate	Anteil schubfreier Patienten	Schubcharakteristika	Anteil Beh.progr.- freier Patienten	EDSS-Änderung	MSFC-Änderung	EQ-5D-Änderung	
FREEDOMS	ja	ja	ja	ja	ja	ja	ja	ja
TRANSFORMS	ja	ja	ja	ja	ja	ja	ja	ja

4.3.1.3.1 Endpunkte – RCT

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

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

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

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und 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). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Im Folgenden sind die Endpunkte Schübe (Operationalisierungen Jährliche Schubrate, Anteil schubfreier Patienten, Schubcharakteristika), Behinderungsprogression (Operationalisierung Anteil behinderungsprogressionsfreier Patienten, EDSS-Änderung, MSFC-Änderung), gesundheitsbezogene Lebensqualität (Operationalisierung EQ-5D Änderung im Index und VAS) und unerwünschte Ereignisse (Operationalisierungen Anzahl Patienten mit UE, SUE, Behandlungsabbrüchen aufgrund von UE und mit relevanten UE) dargestellt.

4.3.1.3.1.1 Schübe

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

Tabelle 4-13: Operationalisierung von „Schübe“^{ca}

Studie	Operationalisierung
FREEDOMS	Aggregierte jährliche Schubrate: Gesamtzahl aller bestätigten Schübe aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25
	Anteil schubfreier Patienten zu Monat 24 (Kaplan-Meier Schätzer aus „Zeit bis zum ersten bestätigten Schub“) und daraus ermitteltes Schubrisiko zu Monat 24
	Schubcharakteristika: Anzahl der milden, mittelschweren und schweren Schübe ^b ; Anzahl der Schübe, die mit Steroid behandelt wurden; Anzahl der Schübe, die eine Hospitalisierung erforderten
TRANSFORMS	Aggregierte jährliche Schubrate: Gesamtzahl aller bestätigten Schübe aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25
	Anteil schubfreier Patienten zu Monat 12 (Kaplan-Meier Schätzer aus „Zeit bis zum ersten bestätigten Schub“) und daraus ermitteltes Schubrisiko zu Monat 12
	Schubcharakteristika: Anzahl der milden, mittelschweren und schweren Schübe ^b ; Anzahl der Schübe, die mit Steroid behandelt wurden; Anzahl der Schübe, die eine Hospitalisierung erforderten
<p>^aEin Schub ist allgemein definiert als neue, verschlechterte oder erneut auftretende neurologische Symptome, die mindestens 30 Tage nach Beginn des vorausgegangenen Schubs auftraten, mindestens 24 Stunden anhielten, ohne Fieber oder Infektion verliefen.</p> <p>Ein bestätigter Schub liegt vor, wenn die allgemeinen Symptome von einem Anstieg um $\geq 0,5$ EDSS-Punkte oder um 1 Punkt in zwei Funktionssystem-Werten oder um 2 Punkte in einem Funktionssystem-Wert (mit Ausnahme von Darm/Blase und psychischen Systemen) begleitet waren.</p> <p>b: ein milder Schub ist definiert als EDSS-Anstieg um 0,5 Punkte oder 1 Punkt in 1 bis 3 Funktionswerten; ein mittelschwerer Schub ist definiert als EDSS-Anstieg um 1 oder 2 Punkte, eine Änderung um 2 Punkte in 1 oder 2 Funktionswerten oder 1 Punkt Änderung in 4 oder mehr Funktionswerten; ein schwerer Schub ist definiert als ein Schub, der die Kriterien eines mittelschweren Schubes übersteigt.</p>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-14: Bewertung des Verzerrungspotenzials für „Schübe“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
Aggregierte jährliche Schubrate						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
Anteil schubfreier Patienten						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
Schubcharakteristika: Schubschwere						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
Schubcharakteristika: Schübe, die einen Krankenhausaufenthalt erfordern						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
Schubcharakteristika: Schübe, die eine Steroidbehandlung erfordern						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial für die schubbezogenen Endpunkte (alle Operationalisierungen) für beide Studien als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in beiden Fällen als „niedrig“ eingestuft wurde, jeweils der EDSS-Rater (unabhängiger beurteilender Arzt) verblindet war, das ITT-Prinzip adäquat umgesetzt wurde und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Faktoren gab.

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 4-15: Ergebnisse für „Jährliche Schubrate“ aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS	Fingolimod 0,5 mg (n=425)	Placebo (n=418)
ITT ^a	n=425	n=418
Jährliche Schubrate (95 %-KI) über 24 Monate	0,184 (0,152; 0,223)	0,398 (0,339; 0,467)
ARR ratio geg. Placebo (95 %-KI) ^e Vergleich zu Placebo	0,462 (0,373; 0,571) p<0,001	
IFN Non-Responder Def I ^{b,c}	n=57	n=52
Jährliche Schubrate (95 %-KI) über 24 Monate	0,239 (0,152; 0,376)	0,631 (0,439; 0,907)
ARR ratio geg. Placebo (95 %-KI) ^e Vergleich zu Placebo	0,379 (0,212; 0,677) p=0,001	
IFN Non-Responder Def II ^{b,c}	n=60	n=54
Jährliche Schubrate (95 %-KI) über 24 Monate	0,189 (0,117; 0,307)	0,659 (0,461; 0,942)
ARR ratio geg. Placebo (95 %-KI) ^e Vergleich zu Placebo	0,287 (0,158; 0,524) p<0,001	
Patienten mit schnell fortschreitender MS ^{b,d}	n=48	n=37
Jährliche Schubrate (95 %-KI) über 24 Monate	0,243 (0,150; 0,395)	0,739 (0,491; 1,111)
ARR ratio geg. Placebo (95 %-KI) ^e Vergleich zu Placebo	0,329 (0,175; 0,620) p<0,001	
TRANSFORMS	Fingolimod 0,5 mg (n=429)	IFN-β-1a i. m. (n=431)
ITT ^a	n=429	n=431
Jährliche Schubrate (95 %-KI) über 12 Monate	0,161 (0,122; 0,212)	0,331 (0,262; 0,417)
ARR ratio geg. IFN-β-1a i. m. (95 %- KI) ^e Vergleich zu IFN-β-1a i. m.	0,484 (0,367; 0,639) p<0,001	
IFN Non-Responder Def I ^{b,c}	n=160	n=149
Jährliche Schubrate (95 %-KI) über 12 Monate	0,201 (0,138; 0,293)	0,514 (0,390; 0,678)
ARR ratio geg. IFN-β-1a i. m. (95 %- KI) ^e	0,391 (0,246; 0,623)	

Vergleich zu IFN- β -1a i. m.	p<0,001	
IFN Non-Responder Def II ^{b,c}	n=166	n=138
Jährliche Schubrate (95 %-KI) über 12 Monate	0,200 (0,138; 0,289)	0,514 (0,387; 0,684)
ARR ratio geg. IFN- β -1a i. m. (95 %-KI) ^e	0,389 (0,244; 0,619)	
Vergleich zu IFN- β -1a i. m.	p<0,001	
Patienten mit schnell fortschreitender MS ^{b, d}	n=27	n=30
Jährliche Schubrate (95 %-KI) über 12 Monate	0,226 (0,094; 0,542)	0,303 (0,146; 0,631)
ARR ratio geg. IFN- β -1a i. m. (95 %-KI) ^e	0,746 (0,238; 2,333)	
Vergleich zu IFN- β -1a i. m.	p=0,614	
<p>a: Effektschätzer und <i>p</i>-Werte für Fingolimod gegenüber den Vergleichsbehandlungen wurden mithilfe des negativen Binomialregressionsmodells berechnet, nach Behandlungsgruppe, Land, Schubzahl in den 2 Jahren vor Studienbeginn und dem EDSS-Wert (Expanded Disability Status Scale) zu Studienbeginn adjustiert</p> <p>b: Effektschätzer und <i>p</i>-Werte für Fingolimod gegenüber den Vergleichsbehandlungen wurden mithilfe des negativen Binomialregressionsmodells berechnet, nach Behandlungsgruppe adjustiert (Anwendung eines reduzierten Modells, da das Modell mit 4 Covariaten wie unter a genannt nicht konvergiert)</p> <p>c: Anwendungsgebiet a1 und a2, siehe Modul 3A Abschnitt 3.1.2</p> <p>d: Anwendungsgebiet b, siehe Modul 3A Abschnitt 3.1.2</p> <p>e: Bei einem Assoziationsmaß wie der relativen Rate ist es von Interesse zu beurteilen, ob das relative Risiko größer oder kleiner als 1 ist. Um dies beurteilen zu können, ist die Angabe eines Konfidenzintervalls an Stelle des Standardfehlers sinnvoll.</p>		

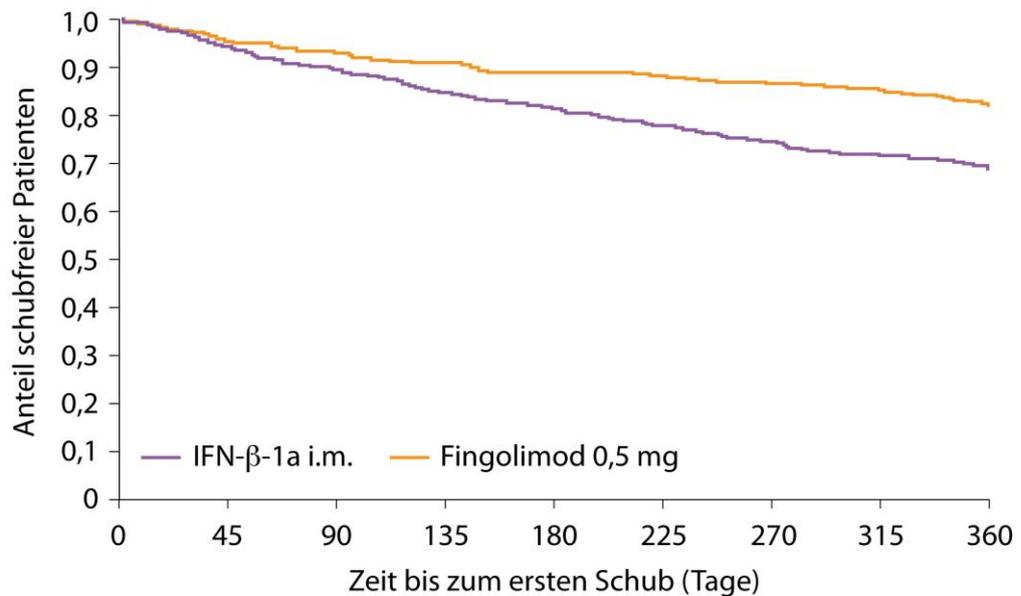
In beiden Phase III-Studien senkte Fingolimod 0,5 mg die jährliche Schubrate (ARR) signifikant verglichen mit den Vergleichsbehandlungen (Placebo und zweckmäßige Vergleichstherapie IFN- β -1a i. m.) und erfüllte erfolgreich die primären Ziele der Phase III-Hauptstudien. Fingolimod 0,5 mg senkte die jährliche Schubrate in FREEDOMS nach 2 Jahren der Behandlung signifikant um 54 % im Vergleich zu Placebo (ARR ratio 0,462; 95 %-KI (0,373; 0,571); $p < 0,001$; Tabelle 4-15). In TRANSFORMS senkte Fingolimod 0,5 mg die jährliche Schubrate nach 1 Jahr der Behandlung im Vergleich zu IFN- β -1a i. m. um 52 % (ARR ratio 0,484; 95 %-KI (0,367; 0,639); $p < 0,001$; Tabelle 4-15). In den Auswertungen zur Zielpopulation zeigt sich gegenüber Placebo in allen drei definierten Subpopulationen ein konsistenter signifikanter Effekt mit einer Senkung der jährlichen Schubrate zwischen 62 % und 71 % (ARR ratios und 95 %-KIs für IFN Non-Responder Def I und II und Patienten mit schnell fortschreitender schwerer RRMS: 0,379 (0,212; 0,677), 0,287 (0,158; 0,524), 0,329 (0,175; 0,620)). Für die beiden Subpopulationen der IFN Non-Responder verhält es sich gegenüber der zweckmäßige Vergleichstherapie IFN- β -1a i. m. ebenso. Die jährliche Schubrate wird dabei signifikant um 61 % gegenüber IFN- β -1a i. m. gesenkt. Die Senkung der jährlichen Schubrate um 25 % bei den Patienten mit schnell fortschreitender MS war hingegen nicht signifikant, die Gruppengröße mit rund 30 Patienten aber auch nicht aussagekräftig (ARR ratios und 95 %-KIs für IFN Non-Responder Def I und II und Patienten mit schnell fortschreitender schwerer RRMS: 0,391 (0,246; 0,623), 0,389 (0,244; 0,619), 0,746 (0,238; 2,333)).

Tabelle 4-16: Ergebnisse für „Anteil schubfreier Patienten“ aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS	Fingolimod 0,5 mg (n=425)	Placebo (n=418)
ITT ^a	n=425	n=418
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	70,4 (2,26)	45,6 (2,52)
95 %-KI des KM-Schätzers	(65,95; 74,80)	(40,70; 50,57)
Hazard Ratio (95 %-KI) ^e	0,48 (0,39; 0,61)	
Vergleich gegen Placebo	p<0,001	
IFN Non-Responder Def I ^{b,c}	n=57	n=52
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	68,3 (6,38)	38,2 (7,15)
95 %-KI des KM-Schätzers	(53,97; 78,96)	(24,40; 51,78)
Hazard Ratio (95 %-KI) ^e	0,40 (0,22; 0,73)	
Vergleich gegen Placebo	p=0,003	
IFN Non-Responder Def II ^{b,c}	n=60	n=54
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	75,3 (5,74)	32,7 (7,03)
95 %-KI des KM-Schätzers	(61,85; 84,58)	(19,59; 46,45)
Hazard Ratio (95 %-KI) ^e	0,29 (0,15; 0,53)	
Vergleich gegen Placebo	p<0,001	
Patienten mit schnell fortschreitender MS ^{b,d}	n=48	n=37
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	70,1 (6,70)	26,3 (7,47)
95 %-KI des KM-Schätzers	(54,81; 81,13)	(13,14; 41,43)
Hazard Ratio (95 %-KI) ^e	0,32 (0,16; 0,61)	
Vergleich gegen Placebo	p<0,001	
TRANSFORMS^a	Fingolimod 0,5 mg (n=429)	IFN-β-1a i. m. (n=431)
ITT ^a	n=429	n=431
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	82,6 (1,85)	69,3 (2,28)
95 %-KI des KM-Schätzers	(79,02; 86,27)	(64,84; 73,76)
Hazard Ratio (95 %-KI) ^e	0,52 (0,39; 0,69)	
Vergleich gegen IFN-β-1a i. m.	p<0,001	
IFN Non-Responder Def I ^{b,c}	n=160	n=149
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	83,6 (2,95)	61,6 (4,13)
95 %-KI des KM-Schätzers	(76,81; 88,50)	(52,99; 69,13)
Hazard Ratio (95 %-KI) ^e	0,38 (0,24; 0,61)	

Vergleich gegen IFN- β -1a i. m. ^a	p<0,001	
IFN Non-Responder Def II ^{b,c}	n=166	n=138
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	83,5 (2,90)	64,2 (4,20)
95 %-KI des KM-Schätzers	(76,93; 88,41)	(55,37; 71,81)
Hazard Ratio (95 %-KI) ^e	0,42 (0,26; 0,67)	
Vergleich gegen IFN- β -1a i. m.	p<0,001	
Patienten mit schnell fortschreitender MS ^{b,d}	n=27	n=30
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	80,1 (7,99)	76,3 (7,85)
95 %-KI des KM-Schätzers	(58,54; 91,25)	(56,54; 87,92)
Hazard Ratio (95 %-KI) ^e	0,76 (0,24; 2,39)	
Vergleich gegen IFN- β -1a i. m.	p=0,637	
<p>a: HRs und p-Werte wurden mittels des proportionalen Risikomodells nach Cox berechnet, nach Behandlungsgruppe, Land, Schubzahl in den 2 Jahren vor Studienbeginn und dem EDSS-Wert (Expanded Disability Status Scale) zu Studienbeginn adjustiert.</p> <p>b: HRs und p-Werte wurden mittels des proportionalen Risikomodells nach Cox berechnet und nach Behandlungsgruppe adjustiert (Anwendung eines reduzierten Modells, da das Modell mit 4 Kovariaten wie unter a genannt nicht konvergiert).</p> <p>c: Anwendungsgebiet a1 und a2, siehe Modul 3A Abschnitt 3.1.2.</p> <p>d: Anwendungsgebiet b, siehe Modul 3A Abschnitt 3.1.2.</p> <p>e: Bei einem Assoziationsmaß wie dem relativen Risiko ist es von Interesse zu beurteilen, ob das relative Risiko größer oder kleiner als 1 ist. Um dies beurteilen zu können, ist die Angabe eines Konfidenzintervalls sinnvoll an Stelle des Standardfehlers.</p>		

In beiden Phase III-Studien wurde der Anteil schubfreier Patienten in den Fingolimod-Behandlungsgruppen signifikant erhöht gegenüber den Vergleichsbehandlungen. In FREEDOMS waren 70 % (95 %-KI: (65,95; 74,80)) der mit Fingolimod 0,5 mg behandelten Patienten nach 24 Monaten schubfrei im Vergleich zu 46 % (95 %-KI: (40,70; 50,57)) unter Placebo. In TRANSFORMS waren 83 % (95 %-KI: (79,02; 86,27)) der mit Fingolimod 0,5 mg behandelten Patienten nach 12 Monaten schubfrei im Vergleich zu 69 % (95 %-KI: (64,84; 73,76)) der mit IFN- β -1a i. m. behandelten Patienten. Schon im Verlauf zeigte sich, dass unter Fingolimod 0,5 mg mehr Patienten schubfrei waren als unter IFN- β -1a i. m. (Abbildung 4-17). Zensierungen im Verlauf begründen sich auf Studienabbrüche (Darstellungen inkl. Zensierungen finden sich im Anhang 4-I).



Anzahl Patienten mit Risiko für den ersten Schub (number at risk)

Fingolimod 0,5 mg	429	408	397	385	374	368	359	351	294
IFN-β-1a i.m.	431	402	379	350	335	313	299	286	240

Abbildung 4-17: Anteil schubfreier Patienten unter Fingolimod verglichen mit der zweckmäßigen Vergleichstherapie IFN-β-1a i. m. in TRANSFORMS (ITT-Population) Intent-to-treat-Population. Daten werden als Prozentsatz von Patienten ohne bestätigten Schub basierend auf Kaplan-Meier-Schätzungen dargestellt.

In den der Zielpopulation entsprechenden Subpopulationen sind die Daten konsistent und zeigen ebenfalls einen signifikanten Vorteil gegenüber Placebo beziehungsweise gegenüber der zweckmäßigen Vergleichstherapie IFN-β-1a i. m., die Größe des Effekts spricht in den Subpopulationen (abgesehen von den Patienten mit schnell fortschreitender schwerer MS in der TRANSFORMS-Studie) sogar noch deutlicher für Fingolimod im Vergleich zu den jeweiligen Komparatoren. Lediglich bei den Patienten mit schnell fortschreitender MS war der Effekt nicht signifikant unterschiedlich, die Gruppengröße mit rund 30 Patienten dafür aber auch nicht aussagekräftig. Kaplan-Meier-Darstellungen inklusive Zensierungen finden sich im Anhang 4-I, Zensierungen im Verlauf begründen sich auf Studienabbrüche.

Die Daten unterstützen nachdrücklich den klinischen Zusatznutzen von Fingolimod 0,5 mg gegenüber der zweckmäßigen Vergleichstherapie IFN-β-1a i. m. für RRMS und zeigen eine deutliche Wirkung auf die Schubrate.

Tabelle 4-17: Ergebnisse für „Schubcharakteristika“ aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS^{a,b}	Fingolimod 0,5 mg (n=425)	Placebo (n=418)
ITT ^{a,b}	n=425	n=418
Schubzahl gesamt, n	172	359
Nach Schubschwere, n(%)		
Mild	60 (34,9)	119 (33,1)
Mittelschwer	102 (59,3)	205 (57,1)
Schwer	10 (5,8)	35 (9,7)
Vergleich gegen Placebo	p=0,322 (overall)	
Davon Steroid-behandelt, n(%)	140 (81,4)	303 (84,4)
Vergleich gegen Placebo	p=0,385	
Davon hospitalisierungspflichtig, n(%)	63 (36,6)	146 (40,7)
Vergleich gegen Placebo	p=0,394	
IFN Non-Responder Def I ^{a,b,c}	n=57	n=52
Schubzahl gesamt, n	25	54
Nach Schubschwere, n(%)		
Mild	10 (40,0)	22 (40,7)
Mittelschwer	14 (56,0)	25 (46,3)
Schwer	1 (4,0)	7 (13,0)
Vergleich gegen Placebo	p=0,538 (overall)	
Davon Steroid-behandelt, n(%)	16 (64,0)	45 (83,3)
Vergleich gegen Placebo	p=0,083	
Davon hospitalisierungspflichtig, n(%)	8 (32,0)	17 (31,5)
Vergleich gegen Placebo	p=1,000	
IFN Non-Responder Def II ^{a,b,c}	n=60	n=54
Schubzahl gesamt, n	21	56
Nach Schubschwere, n(%)		
Mild	9 (42,9)	20 (35,7)
Mittelschwer	11 (52,4)	29 (51,8)
Schwer	1 (4,8)	7 (12,5)
Vergleich gegen Placebo	p=0,660 (overall)	
Davon Steroid-behandelt, n(%)	14 (66,7)	49 (87,5)
Vergleich gegen Placebo	p=0,048	

Davon hospitalisierungspflichtig, n(%) Vergleich gegen Placebo	8 (38,1) p=0,589	17 (30,4)
Patienten mit schnell fortschreitender MS ^{a,b,d}	n=48	n=37
Schubzahl gesamt, n	22	46
Nach Schubschwere, n(%)		
Mild	4 (18,2)	10 (21,7)
Mittelschwer	17 (77,3)	32 (69,6)
Schwer	1 (4,5)	4 (8,7)
Vergleich gegen Placebo	p=0,833 (overall)	
Davon Steroid-behandelt, n(%) Vergleich gegen Placebo	21 (95,5) p=0,149	37 (80,4)
Davon hospitalisierungspflichtig, n(%) Vergleich gegen Placebo	6 (27,3) p=0,296	19 (41,3)
TRANSFORMS^a	Fingolimod 0,5 mg (n=429)	IFN-β-1a i. m. (n=431)
ITT ^{a,b}	n=429	n=431
Schubzahl gesamt, n	89	179
Nach Schubschwere, n(%)		
Mild	31 (34,8)	44 (24,6)
Mittelschwer	48 (53,9)	105 (58,7)
Schwer	10 (11,2)	30 (16,8)
Vergleich gegen IFN-β-1a i. m.	p=0,166(overall)	
Davon Steroid-behandelt, n(%) Vergleich gegen IFN-β-1a i. m.	65 (73,0) p=0,078	148 (82,7)
Davon hospitalisierungspflichtig, n(%) Vergleich gegen IFN-β-1a i. m.	11 (12,4) p=0,128	36 (20,1)
IFN Non-Responder Def I ^{a,b,c}	n=160	n=149
Schubzahl gesamt, n	32	72
Nach Schubschwere, n(%)		
Mild	6 (18,8)	18 (25,0)
Mittelschwer	20 (62,5)	37 (51,4)
Schwer	6 (18,8)	17 (23,6)
Vergleich gegen IFN-β-1a i. m.	p=0,656 (overall)	

Davon Steroid-behandelt, n(%)	26 (81,3)	63 (87,5)
Vergleich gegen IFN- β -1a i. m.	p=0,546	
Davon hospitalisierungspflichtig, n(%)	4 (12,5)	14 (19,4)
Vergleich gegen IFN- β -1a i. m.	p=0,575	
IFN Non-Responder Def II ^{a,b,c}	n=166	n=138
Schubzahl gesamt, n	33	67
Nach Schubschwere, n(%)		
Mild	8 (24,2)	15 (22,4)
Mittelschwer	20 (60,6)	35 (52,2)
Schwer	5 (15,2)	17 (25,4)
Vergleich gegen IFN- β -1a i. m.	p=0,552 (overall)	
Davon Steroid-behandelt, n(%)	24 (72,7)	57 (85,1)
Vergleich gegen IFN- β -1a i. m.	p=0,177	
Davon hospitalisierungspflichtig, n(%)	4 (12,1)	11 (16,4)
Vergleich gegen IFN- β -1a i. m.	p=0,768	
Patienten mit schnell fortschreitender MS ^{a,b,d}	n=27	n=30
Schubzahl gesamt, n	6	9
Nach Schubschwere, n(%)		
Mild	2 (33,3)	3 (33,3)
Mittelschwer	4 (66,7)	5 (55,6)
Schwer	0 (0,0)	1 (11,1)
Vergleich gegen IFN- β -1a i. m.	p=1,000 (overall)	
Davon Steroid-behandelt, n(%)	3 (50,0)	6 (66,7)
Vergleich gegen IFN- β -1a i. m.	p=0,622	
Davon hospitalisierungspflichtig, n(%)	0 (0,0)	1 (11,1)
Vergleich gegen IFN- β -1a i. m.	p=1,000	
a: Prozentangaben beziehen sich jeweils auf die gesamte Schubzahl		
b: p-Werte sind mittels Fisher's Exact Test berechnet		
c: Anwendungsgebiet a1 und a2, siehe Modul 3A Abschnitt 3.1.2		
d: Anwendungsgebiet b, siehe Modul 3A Abschnitt 3.1.2		
Anmerkung: keine Angabe von Standardfehler/konfidenzintervallen; es handelt sich um die Darstellung von Ereignishäufigkeiten, die standardmäßig mit Anzahl und Prozent dargestellt werden		

Die Gesamtzahl der Schübe und konsistent auch die Zahl schwerer Schübe oder Schübe, die einen Krankenhausaufenthalt beziehungsweise eine Steroidbehandlung erfordern, ist bei mit Fingolimod behandelten Patienten gegenüber mit IFN- β -1a i. m. behandelten Patienten reduziert. Diese Schübe sind im Hinblick auf die Kosten, die für das Gesundheitssystem

entstehen, besonders relevant. Die Kosten, die durch einen hospitalisierungspflichtigen Schub entstehen, können bis um den Faktor 50 über jenen eines milden Schubes liegen wenn die Kosten für Neuro-Rehabilitationsmaßnahmen einbezogen werden (2). Selbst ohne diese Aufwendungen können die Kosten einen Wert von ungefähr 2.300 EUR pro Schub erreichen, bei Notwendigkeit einer Plasmapherese sogar bis zu rund 9.000 EUR¹³. Eine Reduktion der Anzahl hospitalisierungspflichtiger Schübe ist daher besonders bedeutsam. Die Reduktion der Anzahl schwerer Schübe und hospitalisierungspflichtiger Schübe war in der Gesamtpopulation sowie den gemäß Label relevanten Subpopulationen zu sehen (Tabelle 4-17). Besonders auffällig ist dabei, dass in der Subpopulation der Patienten mit schnell fortschreitender schwerer RRMS die Anzahl schwerer und hospitalisierungspflichtiger Schübe nicht nur konsistent mit der Gesamtzahl der Schübe sinkt, sondern der Anteil Schübe dieser Schubkategorie sogar prozentual deutlich reduziert wird. Auch wenn also die Schubrate in dieser Subpopulation nur mäßig und nicht signifikant gesenkt werden kann, so zeigt sich doch, dass die Schubschwere deutlich abnimmt und damit ein reduzierender Effekt auf die dadurch entstehenden Kosten zu erwarten ist.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Bei den beiden vorliegenden Phase III-Studien handelt es sich um verschiedene Studientypen (aktiv- beziehungsweise Placebo-kontrolliert), daher erscheint eine Zusammenfassung mittels Meta-Analyse wenig sinnvoll.

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.

¹³ für Patienten mit MS-bedingtem Schub wird folgender ICD-10-Code verwendet: G35.11 ("Multiple Sklerose mit vorherrschend schubförmigem Verlauf: Mit Angabe einer akuten Exazerbation oder Progression" (33)); laut G-DRG Browser der Hauptabteilungen Version 2011 (34) sind die meisten Fälle mit G35.11 in der DRG B68D, diese hat ein Relativgewicht von 0,777 (35); bei einem einheitlichem Basisfallwert für 2011 von 2.971 (36), 16 Euro macht das 2.308,59 Euro für den KH-Aufenthalt; bei Wiendl et al (37) wird auf die Plasmapherese bei besonders schweren Schüben eingegangen, wobei 5 Zyklen empfohlen werden: OPS Version 2011: 8-820.04 ("Therapeutische Plasmapherese" mit normalem Plasma (5 Plasmapheresen)), Fallpauschalenkatalog G-DRG 2011 Abschnitt Zusatzentgelte; hierfür kann das Zusatzentgelt ZE36.05 abgerechnet werden, was 6.527,60 Euro ausmacht. Bezieht man sich nur auf die stationären Behandlungskosten im Krankenhaus bei einem schweren Schub, entstehen somit Kosten i.H.v. 8.836,19 Euro (ohne Reha).

4.3.1.3.1.2 Behinderungsprogression

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

Tabelle 4-18: Operationalisierung von „Behinderungsprogression“

Studie	Operationalisierung
FREEDOMS	Anteil Patienten ohne nach 3 Monaten bestätigter Behinderungsprogression ^a (Kaplan-Meier-Schätzer) zu Monat 24 Mittlere Änderung des EDSS zu Monat 24 gegenüber Baseline Mittlere Änderung des MSFC zu Monat 24 gegenüber Baseline
TRANSFORMS	Anteil Patienten ohne nach 3 Monaten bestätigter Behinderungsprogression ^a (Kaplan-Meier-Schätzer) zu Monat 12 Mittlere Änderung des EDSS zu Monat 12 gegenüber Baseline Mittlere Änderung des MSFC zu Monat 12 gegenüber Baseline
a: Die Behinderungsprogression war definiert als eine Erhöhung des EDSS-Werts um 1 Punkt (um 0,5 Punkte bei einem Baseline-EDSS-Wert $\geq 5,5$), die 3 Monate später ohne Vorliegen eines Schubs bestätigt wurde.	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-19: Bewertung des Verzerrungspotenzials für „Behinderungsprogression“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
Anteil behinderungsprogressionsfreier Patienten						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
Mittlere Änderung des EDSS						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
Mittlere Änderung des MSFC						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial für die Endpunkte zur Behinderung (alle Operationalisierungen) für beide Studien als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in beiden Fällen als „niedrig“ eingestuft wurde, jeweils der EDSS-Rater (unabhängiger beurteilender Arzt) und MSFC-Rater (unabhängiger beurteilender Arzt oder qualifizierte Person) verblindet war, das ITT-Prinzip adäquat umgesetzt wurde und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte gab.

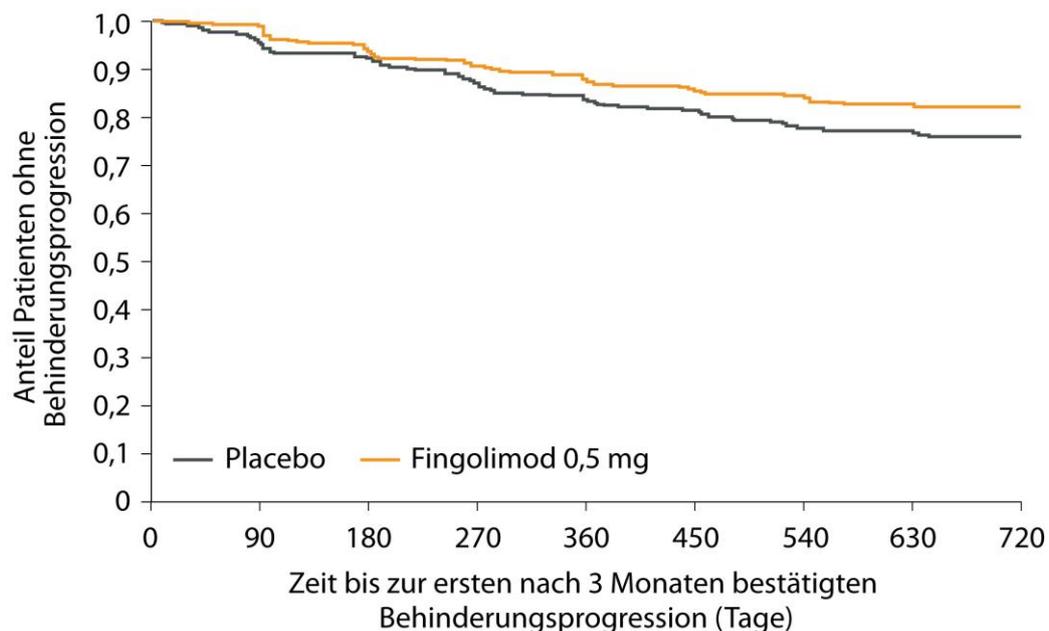
Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-20: Ergebnisse für „Anteil Patienten ohne Behinderungsprogression“ aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS	Fingolimod 0,5 mg N=425	Placebo N=418
ITT^{a,b}	n=425	n=418
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	82,3 (1,89)	75,9 (2,17)
95 %-KI des KM-Schätzers	(78,6; 86,1)	(71,7; 80,2)
Hazard Ratio (95 %-KI) ^e	0,70 (0,52; 0,96)	
Vergleich zu Placebo	p=0,024	
IFN Non-Responder Def I^{a,c,e}	n=57	n=52
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	83,4 (5,05)	75,1 (6,26)
95 %-KI des KM-Schätzers	(70,56; 91,03)	(60,24; 85,05)
Hazard Ratio (95 %-KI) ^e	0,68 (0,29; 1,62)	
Vergleich zu Placebo	p=0,386	
IFN Non-Responder Def II^{a,c,e}	n=60	n=54
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	84,3 (4,82)	74,6 (6,41)
95 %-KI des KM-Schätzers	(71,91; 91,50)	(59,38; 84,76)
Hazard Ratio (95 %-KI) ^e	0,64 (0,27; 1,51)	
Vergleich zu Placebo	p=0,306	
Patienten mit schnell fortschreitender MS^{a,c,f}	n=48	n=37
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	84,7 (5,32)	78,9 (7,12)
95 %-KI des KM-Schätzers	(70,58; 92,41)	(60,63; 89,35)
Hazard Ratio (95 %-KI) ^e	0,73 (0,25; 2,07)	
Vergleich zu Placebo	p=0,548	
TRANSFORMS	Fingolimod 0,5 mg N=429	IFN-β-1a i. m. N=431
ITT^{a,b}	n=429	n=431
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	94,1 (1,15)	92,1 (1,33)
95 %-KI des KM-Schätzers	(91,82; 96,33)	(89,45; 94,66)
Hazard Ratio (95 %-KI)	0,71 (0,42; 1,21)	
Vergleich zu IFN-β-1a i. m.	p=0,209	
IFN Non-Responder Def I^{a,c,e}	n=160	n=149
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	94,3 (1,84)	90,7 (2,47)
95 %-KI des KM-Schätzers	(89,34; 96,99)	(84,47; 94,48)
Hazard Ratio (95 %-KI) ^e	0,60 (0,26; 1,41)	
Vergleich zu IFN-β-1a i. m.	p=0,244	

IFN Non-Responder Def II ^{a,c,e}	n=166	n=138
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	95,1 (1,68)	91,5 (2,45)
95 %-KI des KM-Schätzers	(90,48; 97,53)	(85,20; 95,22)
Hazard Ratio (95 %-KI) ^e	0,58 (0,23; 1,44)	
Vergleich zu IFN-β-1a i. m.	p=0,242	
Patienten mit schnell fortschreitender MS ^{a,c,f}	n=27	n=30
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	100,0	86,5 (6,27)
95 %-KI des KM-Schätzers	(nicht bestimmbar) ^d	(67,99; 94,73)
Hazard Ratio (95 %-KI) ^e	(nicht bestimmbar) ^d	
Vergleich zu IFN-β-1a i. m.	(nicht bestimmbar) ^d	
<p>a: Daten werden als Prozentsatz von Patienten ohne ein nach 3 Monaten bestätigtes Fortschreiten der Behinderung basierend auf Kaplan-Meier-Schätzungen (95 % KI) dargestellt.</p> <p>b: HRs und <i>p</i>-Werte gelten für Fingolimod vs. Placebo oder IFN-β-1a i. m.; Die <i>p</i>-Werte für die HRs wurden mittels des proportionalen Risikomodells nach Cox berechnet, adjustiert nach Behandlung, Land, EDSS-Wert zu Studienbeginn und Alter.</p> <p>c: HRs und <i>p</i>-Werte gelten für Fingolimod vs. Placebo oder IFN-β-1a i. m.; Die <i>p</i>-Werte für die HRs wurden mittels des proportionalen Risikomodells nach Cox berechnet, adjustiert nach Behandlung. (Anwendung eines reduzierten Modells, da das Modell mit 4 Covariaten wie unter b genannt nicht konvergiert)</p> <p>d: kein Fall von Behinderungsprogression</p> <p>e: Anwendungsgebiet a1 und a2, siehe Modul 3A Abschnitt 3.1.2</p> <p>f: Anwendungsgebiet b, siehe Modul 3A Abschnitt 3.1.2</p> <p>e: Bei einem Assoziationsmaß wie der relativen Rate ist es von Interesse zu beurteilen, ob das relative Risiko größer oder kleiner als 1 ist. Um dies beurteilen zu können, ist die Angabe eines Konfidenzintervalls sinnvoll an Stelle des Standardfehlers.</p>		

In FREEDOMS wurde das Risiko einer nach 3 Monaten bestätigten Behinderungsprogression über 2 Jahre unter Fingolimod 0,5 mg im Vergleich zu Placebo um 30 % reduziert (HR: 0,70; (95 %-KI: 0,52; 0,96), *p*=0,024). Ein signifikant größerer Anteil der Patienten blieb unter Fingolimod 0,5 mg frei von einer bestätigten Progression der Behinderung als unter Placebo (siehe Abbildung 4-18). Zensierungen begründen sich auf Studienabbruchern, die hohe Zahl an Zensierungen in den letzten 3 Monaten vor allem auf nicht mehr bestätigten Fälle von Behinderungsprogression. Das sind Fälle, die nach Monat 21 auftraten, deren Termin für eine Bestätigung nach 3 Monaten aber nach Monat 24 gelegen hätten (Darstellungen inkl. Zensierungen finden sich in Anhang 4-I).



Patienten mit Risiko für die erste bestätigte Behinderungsprogression (number at risk)

Fingolimod 0,5 mg	425	416	388	370	354	340	332	321	152
Placebo	418	391	371	341	320	308	290	279	143

Abbildung 4-18: Zeit bis zum Auftreten einer nach 3 Monaten bestätigten Behinderungsprogression in FREEDOMS (ITT-Population)

Intent-to-treat-Population. Daten werden als Prozentsatz von Patienten mit bestätigter Behinderungsprogression basierend auf Kaplan-Meier-Schätzungen dargestellt. Die p -Werte für Fingolimod versus Placebo wurde mithilfe einer Log-Rank-Test-Hauptanalyse und eines proportionalen Risikomodells nach Cox berechnet, adjustiert nach Behandlung, Land, EDSS-Wert zu Studienbeginn und Alter.

Für die gemäß Zielpopulation definierten Subpopulationen war ein konsistenter Trend hin zu einem Vorteil von Fingolimod gegenüber Placebo zu sehen, die Ergebnisse waren jedoch nicht statistisch signifikant. Dies kann in der kleineren Fallzahl gegenüber der Gesamtpopulation und damit der geringen Anzahl an Fällen bestätigter Behinderungsprogression in den einzelnen Subpopulationen begründet sein. Zusammengefasst deuten diese Daten darauf hin, dass Fingolimod übereinstimmende, klinisch bedeutsame Effekte auf die Behinderungsprogression hat. Kaplan-Meier-Darstellungen inklusive Zensierungen finden sich im Anhang 4-I, Zensierungen im Verlauf begründen sich auf Studienabbrüche.

Die Studie TRANSFORMS war aufgrund der Studiendauer von 12 Monaten nicht darauf ausgelegt, einen Unterschied in einer bestätigten Behinderungsprogression zwischen Fingolimod und IFN- β -1a i. m. nachzuweisen. Dennoch sprach die HR für eine bestätigte Progression nach 12 Monaten in TRANSFORMS für Fingolimod 0,5 mg (HR: 0,71; (95 %-KI: 0,42; 1,21), $p=0,209$). Die Trends der Wirkung scheinen mit den Erkenntnissen aus FREEDOMS übereinzustimmen. Dies gilt ebenso für die Subpopulationen gemäß Label.

Tabelle 4-21: Ergebnisse für „Mittlere Änderung des EDSS gegenüber Baseline“ aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS	Fingolimod 0,5 mg (N=425)	Placebo (N=418)
ITT^a	n=374	n=332
Mittlere (SEM) Veränderung EDSS zu Baseline	0,00 (0,045)	0,13 (0,051)
Vergleich zu Placebo	p=0,004	
<u>IFN Non-Responder Def I</u> ^{a,b}	n=48	n=36
Mittlere (SEM) Veränderung EDSS zu Baseline	0,04 (0,142)	-0,15 (0,163)
Vergleich zu Placebo	p=0,852	
<u>IFN Non-Responder Def II</u> ^{a,b}	n=51	n=34
Mittlere (SEM) Veränderung EDSS zu Baseline	-0,03 (0,139)	-0,07 (0,164)
Vergleich zu Placebo	p=0,736	
<u>Patienten mit schnell fortschreitender MS</u> ^{a,c}	n=41	n=26
Mittlere (SEM) Veränderung EDSS zu Baseline	-0,04 (0,148)	-0,04 (0,211)
Vergleich zu Placebo	p=0,818	
TRANSFORMS	Fingolimod 0,5 mg (N=429)	IFN-β-1a i. m. (N=431)
ITT^a	n=394	n=377
Mittlere (SEM) Veränderung EDSS zu Baseline	-0,08 (0,040)	0,01 (0,040)
Vergleich zu IFN-β-1a i. m.	p=0,059	
<u>IFN Non-Responder Def I</u> ^{a,b}	n=149	n=124
Mittlere (SEM) Veränderung EDSS zu Baseline	-0,09 (0,063)	0,04 (0,073)
Vergleich zu IFN-β-1a i. m.	p=0,116	
<u>IFN Non-Responder Def II</u> ^{a,b}	n=154	n=114
Mittlere (SEM) Veränderung EDSS zu Baseline	-0,11 (0,060)	0,04 (0,079)
Vergleich zu IFN-β-1a i. m.	p=0,074	
<u>Patienten mit schnell fortschreitender MS</u> ^{a,c}	n=24	n=27
Mittlere (SEM) Veränderung EDSS zu Baseline	-0,06 (0,063)	0,17 (0,139)
Vergleich zu IFN-β-1a i. m.	p=0,098	
a: Die p-Werte für Fingolimod gegenüber den Vergleichsbehandlungen wurden mithilfe des Wilcoxon-Tests berechnet.		
b: Anwendungsgebiet a1 und a2, siehe Modul 3A Abschnitt 3.1.2.		

c: Anwendungsgebiet b, siehe Modul 3A Abschnitt 3.1.2.

Die Daten der Phase III-Studien zeigen, dass die mittleren EDSS-Werte unter Fingolimod-Behandlung stabil blieben oder sich leicht verbesserten, während sie sich in den Gruppen (ITT-Population) mit Placebo und IFN- β -1a i. m. jedoch leicht verschlechterten (Tabelle 4-21). Sowohl in FREEDOMS als auch in TRANSFORMS sprachen die EDSS-Werte für Fingolimod 0,5 mg im Vergleich zu den jeweiligen Vergleichsbehandlungen.

In den Subpopulationen zeigt sich ein ähnliches Bild. Tendenziell bietet Fingolimod hier einen Vorteil gegenüber den Vergleichsbehandlungen.

Tabelle 4-22: Ergebnisse für „Mittlere Änderung des MSFC gegenüber Baseline“ aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS	Fingolimod 0,5 mg (N=425)	Placebo (N=418)
ITT^a	n=361	n=316
Mittlere (SEM) Veränderung MSFC zu Baseline	0,03 (0,021)	-0,06 (0,032)
Vergleich zu Placebo	p=0,009	
<u>IFN Non-Responder Def I</u> ^{a,b}	n=46	n=33
Mittlere (SEM) Veränderung MSFC zu Baseline	-0,01 (0,055)	-0,24 (0,094)
Vergleich zu Placebo	p=0,144	
<u>IFN Non-Responder Def II</u> ^{a,b}	n=49	n=31
Mittlere (SEM) Veränderung MSFC zu Baseline	-0,05 (0,057)	-0,27 (0,095)
Vergleich zu Placebo	p=0,083	
<u>Patienten mit schnell fortschreitender MS</u> ^{a,c}	n=38	n=25
Mittlere (SEM) Veränderung MSFC zu Baseline	0,07 (0,083)	-0,13 (0,134)
Vergleich zu Placebo	p=0,199	
TRANSFORMS	Fingolimod 0,5 mg (N=429)	IFN-β-1a i. m. (N=431)
ITT^a	n=383	n=366
Mittlere (SEM) Veränderung MSFC zu Baseline	0,04 (0,021)	-0,03 (0,025)
Vergleich zu IFN-β-1a i. m.	p=0,025	
<u>IFN Non-Responder Def I</u> ^{a,b}	n=145	n=124
Mittlere (SEM) Veränderung MSFC zu Baseline	0,02 (0,028)	-0,13 (0,057)
Vergleich zu IFN-β-1a i. m.	p=0,006	
<u>IFN Non-Responder Def II</u> ^{a,b}	n=151	n=114
Mittlere (SEM) Veränderung MSFC zu Baseline	-0,01 (0,030)	-0,09 (0,041)
Vergleich zu IFN-β-1a i. m.	p=0,010	
<u>Patienten mit schnell fortschreitender MS</u> ^{a,c}	n=21	n=26
Mittlere (SEM) Veränderung MSFC zu Baseline	0,02 (0,033)	0,03 (0,061)
Vergleich zu IFN-β-1a i. m.	p=0,856	
a: Die p-Werte für Fingolimod gegenüber den Vergleichsbehandlungen wurden mithilfe des Wilcoxon-Tests berechnet.		
b: Anwendungsgebiet a1 und a2, siehe Modul 3A Abschnitt 3.1.2.		

c: Anwendungsgebiet b, siehe Modul 3A Abschnitt 3.1.2.

Die Daten der Phase III-Studien zeigen, dass die mittleren MSFC z-Punktwerte unter Fingolimod-Behandlung stabil blieben oder sich leicht verbesserten, während sie sich in den Gruppen (ITT-Population) mit Placebo und IFN- β -1a i. m. jedoch leicht verschlechterten (Tabelle 4-22). Sowohl in FREEDOMS als auch in TRANSFORMS sprachen die MSFC z-Punktwerte für Fingolimod 0,5 mg im Vergleich zu den jeweiligen Vergleichsbehandlungen.

In den Subpopulationen zeigt sich ein ähnliches Bild. Tendenziell bietet Fingolimod hier einen Vorteil gegenüber den Vergleichsbehandlungen. Bei den IFN Non-Respondern ist der Vorteil von Fingolimod gegenüber IFN- β -1a i. m. statistisch signifikant. Gegenüber Placebo konnte bei den IFN Non-Respondern kein signifikanter Effekt gesehen werden, was mit der kleinen Fallzahl dieser Subpopulation in der FREEDOMS-Studie erklärt werden kann. Die FREEDOMS-Studie hatte generell einen deutlich kleineren Anteil an vorthera-pierten Patienten und damit auch an Non-Respondern eingeschlossen.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Bei den beiden vorliegenden Phase III-Studien handelt es sich um verschiedene Studientypen (aktiv- beziehungsweise Placebo-kontrolliert), daher erscheint eine Zusammenfassung mittels Meta-Analyse wenig sinnvoll.

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

4.3.1.3.1.3 Gesundheitsbezogene Lebensqualität

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

Tabelle 4-23: Operationalisierung von „gesundheitsbezogener Lebensqualität“

Studie	Operationalisierung
FREEDOMS	Mittlere Änderung des EQ-5D ^a zu Monat 24 gegenüber Baseline
TRANSFORMS	Mittlere Änderung des EQ-5D ^a zu Monat 12 gegenüber Baseline

a: Der Fragebogen EQ-5D ist ein Messinstrument für die gesundheitsbezogene Lebensqualität. Der EQ-5D ist zweigeteilt. Im ersten Teil wird die Lebensqualität über 5 Dimensionen mit jeweils 3 Antwortmöglichkeiten operationalisiert. Die daraus ermittelten Gesundheitszustände werden in einen Lebensqualitätsindex (utility

score) zwischen 0 (Tod) und 1 (bestmöglicher Gesundheitszustand) transformiert. Im zweiten Teil des EQ-5D bewerten die Patienten ihren Gesundheitszustand auf einer visuellen Analogskala (VAS).

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-24: Bewertung des Verzerrungspotenzials für „gesundheitsbezogene Lebensqualität“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
EQ-5D						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „gesundheitsbezogene Lebensqualität“ (EQ-5D) für beide Studien als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in beiden Fällen als „niedrig“ eingestuft wurde, der Patient bei der Bewertung des Gesundheitszustandes verblindet war, das ITT-Prinzip adäquat umgesetzt wurde und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte gab.

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 4-25: Ergebnisse für „gesundheitsbezogene Lebensqualität“ aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS	Fingolimod 0,5 mg (N=425)	Placebo (N=418)
ITT^a		
Mittlere (SEM) Veränderung EQ-5D (Index) zu Baseline	n=366 -0,00 (0,009)	n=326 0,01 (0,011)
Vergleich zu Placebo ^c	p=0,419	
Mittlere (SEM) Veränderung EQ-5D (VAS) zu Baseline	n=365 -0,23 (0,817)	n=326 -1,04 (0,943)
Vergleich zu Placebo	p=0,834	
<u>IFN Non-Responder Def I^{a,b}</u>		
Mittlere (SEM) Veränderung EQ-5D (Index) zu Baseline	n=46 -0,00 (0,032)	n=35 0,05 (0,031)
Vergleich zu Placebo ^c	p=0,523	
Mittlere (SEM) Veränderung EQ-5D (VAS) zu Baseline	n=46 1,02 (2,465)	n=34 2,38 (3,004)
Vergleich zu Placebo	p=0,907	
<u>IFN Non-Responder Def II^{a,b}</u>		
Mittlere (SEM) Veränderung EQ-5D (Index) zu Baseline	n=49 -0,01 (0,031)	n=32 -0,00 (0,020)
Vergleich zu Placebo ^c	p=0,896	
Mittlere (SEM) Veränderung EQ-5D (VAS) zu Baseline	n=49 1,31 (2,786)	n=31 1,00 (3,254)
Vergleich zu Placebo	p=0,730	
<u>Patienten mit schnell fortschreitender MS^{a,c}</u>		
Mittlere (SEM) Veränderung EQ-5D (Index) zu Baseline	n=40 0,03 (0,026)	n=24 -0,07 (0,041)
Vergleich zu Placebo ^c	p=0,037	
Mittlere (SEM) Veränderung EQ-5D (VAS) zu Baseline	n=40 2,08 (2,473)	n=25 -2,96 (4,547)
Vergleich zu Placebo	p=0,251	
TRANSFORMS	Fingolimod 0,5 mg (N=429)	IFN-β-1a i. m. (N=431)
ITT^a		
Mittlere (SEM) Veränderung EQ-5D (Index) zu Baseline	n=374 0,01 (0,009)	n=357 -0,00 (0,010)
Vergleich zu IFN-β-1a i. m.	p=0,100	
Mittlere (SEM) Veränderung EQ-5D	n=375	n=354

(VAS) zu Baseline	1,89 (0,764)	0,63 (0,774)
Vergleich zu IFN- β -1a i. m.	p=0,257	
<u>IFN Non-Responder Def I</u> ^{a,b}		
Mittlere (SEM) Veränderung EQ-5D (Index) zu Baseline	n=138 0,01 (0,015)	n=118 0,03 (0,015)
Vergleich zu IFN- β -1a i. m.	p=0,844	
Mittlere (SEM) Veränderung EQ-5D (VAS) zu Baseline	n=139 3,22 (1,160)	n=118 -0,92 (1,344)
Vergleich zu IFN- β -1a i. m.	p=0,042	
<u>IFN Non-Responder Def II</u> ^{a,b}		
Mittlere (SEM) Veränderung EQ-5D (Index) zu Baseline	n=147 0,01 (0,014)	n=107 0,02 (0,017)
Vergleich zu IFN- β -1a i. m.	p=0,624	
Mittlere (SEM) Veränderung EQ5D (VAS) zu Baseline	n=148 2,95 (1,091)	n=107 -1,40 (1,441)
Vergleich zu IFN- β -1a i. m.	p=0,029	
<u>Patienten mit schnell fortschreitender MS</u> ^{a,c}		
Mittlere (SEM) Veränderung EQ-5D (Index) zu Baseline	n=23 0,04 (0,023)	n=23 -0,02 (0,029)
Vergleich zu IFN- β -1a i. m.	p=0,101	
Mittlere (SEM) Veränderung EQ-5D (VAS) zu Baseline	n=23 2,78 (4,770)	n=22 0,00 (3,006)
Vergleich zu IFN- β -1a i. m.	p=0,699	
a: Die <i>p</i> -Werte für Fingolimod gegenüber den Vergleichsbehandlungen wurden mithilfe des Wilcoxon Tests berechnet.		
b: Anwendungsgebiet a1 und a2, siehe Modul 3A Abschnitt 3.1.2		
c: Anwendungsgebiet b, siehe Modul 3A Abschnitt 3.1.2		

Die mittlere Änderung des aus dem EQ-5D ermittelten gesundheitsbezogenen Lebensqualitätsindex zeigt abgesehen von der Subpopulation der Patienten mit schnell fortschreitender MS keinen Unterschied zwischen Fingolimod und Placebo beziehungsweise IFN- β -1a i. m., der zweckmäßigen Vergleichstherapie. Für die Subpopulation der Patienten mit schnell fortschreitender MS war die mittlere Änderung des Lebensqualitätsindex signifikant unterschiedlich für Fingolimod im Vergleich zu Placebo, wobei sich Fingolimod-Patienten in ihrer gesundheitsbezogenen Lebensqualität verbesserten und Placebo-Patienten verschlechterten. Tendenziell ähnlich fiel auch der Vergleich Fingolimod zu IFN- β -1a i. m. aus, allerdings war der Unterschied hier nicht signifikant.

Die Bewertung der gesundheitsbezogenen Lebensqualität durch den Patienten mittels der VAS zeigte einen nahezu konsistenten Trend zugunsten von Fingolimod. Die Änderung gegenüber Baseline zeigte dabei zumeist eine geringere Verschlechterung oder sogar ausgeprägtere Verbesserung des Zustandes unter Fingolimod im Vergleich zu Placebo oder IFN- β -1a i. m.. Für die Population der IFN Non-Responder war diese Verbesserung statistisch signifikant für Fingolimod im Vergleich zu IFN- β -1a i. m.. Fingolimod scheint hinsichtlich

der Beeinflussung der gesundheitsbezogenen Lebensqualität von Vorteil gegenüber der zweckmäßigen Vergleichstherapie IFN- β -1a i. m. zu sein, mit einem besonders deutlichen Vorteil für die IFN Non-Responder.

Diese Ergebnisse sind aber aufgrund der geringen Spezifität des EQ-5D (indikationsübergreifendes Instrument) eher von untergeordneter Aussagekraft. Zudem konnte keine Aussage getroffen werden, inwiefern eine orale Therapie im Gegensatz zu einer Injektionstherapie hinsichtlich der Lebensqualität von Vorteil ist. Das Double-Dummy Design der TRANSFORMS-Studie lässt das nicht zu, da jeder Patient sowohl eine Kapsel als auch eine Injektion erhalten hat.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Bei den beiden vorliegenden Phase III-Studien handelt es sich um verschiedene Studientypen (aktiv- beziehungsweise Placebo-kontrolliert), daher erscheint eine Zusammenfassung mittels Meta-Analyse wenig sinnvoll.

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

4.3.1.3.1.4 Unerwünschte Ereignisse

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

Tabelle 4-26: Operationalisierung von „Unerwünschten Ereignissen“

Studie	Operationalisierung
FREEDOMS	Anzahl an Patienten mit unerwünschten Ereignissen in 24 Monaten Anzahl an Patienten mit schwerwiegenden unerwünschten Ereignissen in 24 Monaten Anzahl an Patienten mit relevanten ^a unerwünschten Ereignissen in 24 Monaten Anzahl an Patienten mit unerwünschten Ereignissen in 24 Monaten, die zum Absetzen der Studienmedikation geführt haben
TRANSFORMS	Anzahl an Patienten mit unerwünschten Ereignissen in 12 Monaten Anzahl an Patienten mit schwerwiegenden unerwünschten Ereignissen in 12 Monaten Anzahl an Patienten mit relevanten ^a unerwünschten Ereignissen in 12 Monaten Anzahl an Patienten mit unerwünschten Ereignissen in 12 Monaten, die zum Absetzen der Studienmedikation geführt haben
a: relevante Ereignisse sind: Bradykardie/ AV-Überleitungsstörungen, Infektionen, Hypertension, Leberwerterhöhungen, Makulaödem, Malignitäten und QT-Zeit-Verlängerung	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-27: Bewertung des Verzerrungspotenzials für „Unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
Anzahl Patienten mit UE						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
Anzahl Patienten mit SUE						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
Anzahl Behandlungsabbrüche aufgrund von UE						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
Relevante UE						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial für die Endpunkte „gesundheitsbezogene unerwünschte Ereignisse“ (alle Operationalisierungen) für beide Studien als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in beiden Fällen als „niedrig“ eingestuft wurde, der Erfasser der unerwünschten Ereignisse (behandelnder Arzt oder Studienschwester) verblindet war, das ITT-Prinzip adäquat umgesetzt wurde und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte gab.

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 4-28: Ergebnisse für „Unerwünschte Ereignisse“ aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)			
	Fingolimod 0,5 mg (n=425) Anzahl (%) Patienten	Placebo (n=418) Anzahl (%) Patienten	Relatives Risiko	95 % KI ^e
<u>Sicherheitspopulation</u>	n=425	n=418		
UE	401 (94,4)	387 (92,6)	1,02	(0,98; 1,06)
schwerwiegende UE	42 (9,9)	53 (12,7)	0,78	(0,53; 1,14)
Behandlungsabbrüche aufgrund von UE	32 (7,5)	32 (7,4)	1,02	(0,63; 1,63)
Relevante UE				
Bradykardie/ AV- Überleitungsstörungen	22 (5,2)	11 (2,6)	1,97	(0,97; 4,01)
Infektion	304 (71,5)	301 (72,0)	0,99	(0,91; 1,08)
Hypertonie	36 (8,5)	22 (5,3)	1,61	(0,96; 2,69)
Leberwerterhöhung	75 (17,6)	27 (6,5)	2,73	(1,80; 4,15)
Lymphopenie	12 (2,8)	2 (0,5)	5,90	(1,33; 26,21)
Leukopenie	10 (2,4)	1 (0,2)	9,84	(1,26; 76,49)
Makulaödem	0	0	-	-
Malignität	5 (1,2)	10 (2,4)	0,49	(0,17; 1,43)
QT-Zeit-Verlängerung	5 (1,2)	5 (1,2)	0,98	(0,29; 3,37)
<u>IFN Non-Responder Def I^c</u>	n=57	n=52		
UE	55 (96,5)	49 (94,2)	1,02	(0,94; 1,11)
schwerwiegende UE	7 (12,3)	10 (19,2)	0,64	(0,26; 1,55)
Behandlungsabbrüche aufgrund von UE	6 (10,5)	5 (9,6)	1,09	(0,36; 3,37)
Relevante UE				
Bradykardie/ AV- Überleitungsstörungen	3 (5,3)	4 (7,7)	0,68	(0,16; 2,91)
Infektion	43 (75,4)	38 (73,1)	1,03	(0,83; 1,29)
Hypertonie	1 (1,8)	2 (3,8)	0,46	(0,04; 4,88)
Leberwerterhöhung	5 (8,8)	4 (7,7)	1,14	(0,32; 4,02)
Lymphopenie	4 (7,0)	0	-	-
Leukopenie	1 (1,8)	1 (1,9)	0,91	(0,06; 14,22)
Makulaödem	0	0	-	-

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

Malignität	1 (1,8)	3 (5,8)	0,30	(0,03; 2,83)
QT-Zeit-Verlängerung	1 (1,8)	1 (1,9)	0,91	(0,06; 14,22)
<u>IFN Non-Responder Def II^c</u>	n=60	n=54		
UE	58 (96,7)	50 (92,6)	1,04	(0,96; 1,14)
schwerwiegende UE	8 (13,3)	10 (18,5)	0,72	(0,31; 1,69)
Behandlungsabbrüche aufgrund von UE	6 (10,0)	4 (7,4)	1,35	(0,40; 4,53)
Relevante UE				
Bradykardie/ AV-Überleitungsstörungen	4 (6,7)	3 (5,6)	1,20	(0,28; 5,12)
Infektion	42 (70,0)	41 (75,9)	0,92	(0,74; 1,15)
Hypertonie	3 (5,0)	2 (3,7)	1,35	(0,23; 7,78)
Leberwerterhöhung	5 (8,3)	3 (5,6)	1,50	(0,38; 5,98)
Lymphopenie	4 (6,7)	0	-	-
Leukopenie	1 (1,7)	1 (1,9)	0,90	(0,06; 14,04)
Makulaödem	0	0	-	-
Malignität	1 (1,7)	2 (3,7)	0,45	(0,04; 4,82)
QT-Zeit-Verlängerung	1 (1,7)	1 (1,9)	0,90	(0,06; 14,04)
<u>Patienten mit schnell fortschreitender MS^d</u>	n=48	n=37		
UE	46 (95,8)	34 (91,9)	1,04	(0,93; 1,17)
schwerwiegende UE	5 (10,4)	3 (8,1)	1,28	(0,33; 5,03)
Behandlungsabbrüche aufgrund von UE	3 (6,3)	1 (2,7)	2,31	(0,25; 21,34)
Relevante UE				
Bradykardie/ AV-Überleitungsstörungen	2 (4,2)	1 (2,7)	1,54	(0,15; 16,36)
Infektion	35 (72,9)	28 (75,7)	0,96	(0,75; 1,24)
Hypertonie	3 (6,3)	1 (2,7)	2,31	(0,25; 21,34)
Leberwerterhöhung	7 (14,6)	4 (10,8)	1,35	(0,43; 4,27)
Lymphopenie	0	1 (2,7)	-	-
Leukopenie	1 (2,1)	0	-	-
Makulaödem	0	0	-	-
Malignität	1 (2,1)	0	-	-
QT-Zeit-Verlängerung	0	1 (2,7)	-	-

TRANSFORMS^b	Fingolimod 0,5 mg (n=429) Anzahl (%) Patienten	IFN-β-1a i. m. (n=431) Anzahl (%) Patienten	Relatives Risiko	95 %-KI ^e
<u>Sicherheitspopulation</u>	n=429	n=431		
UE	369 (86,0)	395 (91,6)	0,94	(0,89; 0,98)
schwerwiegende UE	30 (7,0)	24 (5,6)	1,26	(0,75; 2,11)
Behandlungsabbrüche aufgrund von UE	24 (5,6)	16 (3,7)	1,51	(0,81; 2,80)
Relevante UE				
Bradykardie/ AV- Überleitungsstörungen	15 (3,5)	11 (2,6)	1,37	(0,64; 2,95)
Infektion	222 (51,7)	219 (50,8)	1,02	(0,89; 1,16)
Hypertonie	20 (4,7)	11 (2,6)	1,83	(0,89; 3,77)
Leberwerterhöhung	61 (14,2)	17 (3,9)	3,60	(2,14; 6,07)
Lymphopenie	1 (0,2)	0	-	-
Leukopenie	7 (1,6)	1 (0,2)	7,03	(0,87; 56,92)
Makulaödem	2 (0,5)	1 (0,2)	2,01	(0,18; 22,08)
Malignität	12 (2,8)	3 (0,7)	4,02	(1,14; 14,14)
QT-Zeit-Verlängerung	7 (1,6)	4 (0,9)	1,76	(0,52; 5,96)
grippeartige Erkrankung	15 (3,5)	159 (36,9)	0,09	(0,06; 0,16)
<u>IFN Non-Responder Def I^c</u>	n=160	n=149		
UE	138 (86,3)	133 (89,3)	0,97	(0,89; 1,05)
schwerwiegende UE	13 (8,1)	8 (5,4)	1,51	(0,65; 3,55)
Behandlungsabbrüche aufgrund von UE	7 (4,4)	5 (3,4)	1,30	(0,42; 4,02)
Relevante UE				
Bradykardie/ AV- Überleitungsstörungen	5 (3,1)	3 (2,0)	1,55	(0,38; 6,38)
Infektion	86 (53,8)	81 (54,4)	0,99	(0,80; 1,21)
Hypertonie	10 (6,3)	4 (2,7)	2,33	(0,75; 7,26)
Leberwerterhöhung	17 (10,6)	8 (5,4)	1,98	(0,88; 4,45)
Lymphopenie	1 (0,6)	0	-	-
Leukopenie	6 (3,8)	0	-	-
Makulaödem	1 (0,6)	0	-	-
Malignität	2 (1,3)	0	-	-
QT-Zeit-Verlängerung	2 (1,3)	2 (1,3)	0,93	(0,13; 6,53)
grippeartige Erkrankung	6 (3,8)	41 (27,5)	0,14	(0,06; 0,31)

<u>IFN Non-Responder Def II^c</u>	n=166	n=138		
UE	146 (88,0)	123 (89,1)	0,99	(0,91; 1,07)
schwerwiegende UE	14 (8,4)	7 (5,1)	1,66	(0,69; 4,00)
Behandlungsabbrüche aufgrund von UE	7 (4,2)	4 (2,9)	1,45	(0,43; 4,87)
Relevante UE				
Bradykardie/ AV-Überleitungsstörungen	5 (3,0)	3 (2,2)	1,39	(0,34; 5,69)
Infektion	91 (54,8)	73 (52,9)	1,04	(0,84; 1,28)
Hypertonie	10 (6,0)	5 (3,6)	1,66	(0,58; 4,75)
Leberwerterhöhung	20 (12,0)	8 (5,8)	2,08	(0,94; 4,57)
Lymphopenie	1 (0,6)	0	-	-
Leukopenie	5 (3,0)	0	-	-
Makulaödem	2 (1,2)	0	-	-
Malignität	2 (1,2)	1 (0,7)	1,66	(0,15; 18,14)
QT-Zeit-Verlängerung	2 (1,2)	2 (1,4)	0,83	(0,12; 5,83)
grippeartige Erkrankung	6 (3,6)	39 (28,3)	0,13	(0,06; 0,29)
<u>Patienten mit schnell fortschreitender MS^d</u>	n=27	n=30		
UE	23 (85,2)	25 (83,3)	1,02	(0,82; 1,28)
schwerwiegende UE	1 (3,7)	0	-	-
Behandlungsabbrüche aufgrund von UE	1 (3,7)	0	-	-
Relevante UE				
Bradykardie/ AV-Überleitungsstörungen	0	1 (3,3)	-	-
Infektion	14 (51,9)	16 (53,3)	0,97	(0,59; 1,59)
Hypertonie	2 (7,4)	1 (3,3)	2,22	(0,21; 23,15)
Leberwerterhöhung	2 (7,4)	1 (3,3)	2,22	(0,21; 23,15)
Lymphopenie	0	0	-	-
Leukopenie	0	0	-	-
Makulaödem	0	0	-	-
Malignität	1 (3,7)	0	-	-
QT-Zeit-Verlängerung	0	1 (3,3)	-	-
grippeartige Erkrankung	1 (3,7)	9 (30,0)	0,12	(0,02; 0,91)
a: Unter Fingolimod 0,5 mg traten in FREEDOMS keine Todesfälle auf.				
b: Unter Fingolimod 0,5 mg traten in TRANSFORMS keine Todesfälle auf.				
c: Anwendungsgebiet a1 und a2, siehe Modul 3A Abschnitt 3.1.2.				
d: Anwendungsgebiet b, siehe Modul 3A Abschnitt 3.1.2.				

e: Bei einem Assoziationsmaß wie der relativen Rate ist es von Interesse zu beurteilen, ob das relative Risiko größer oder kleiner als 1 ist. Um dies beurteilen zu können, ist die Angabe eines Konfidenzintervalls sinnvoll an Stelle des Standardfehlers. Anstelle des p-Wertes wurden die Risk Ratios mit den dazugehörigen Konfidenzintervallen angegeben, die die Bewertung der Signifikanz ebenfalls ermöglichen

Dank umfangreicher Studien im Bereich RRMS ist das Sicherheitsprofil von Fingolimod gut beschrieben. In FREEDOMS und TRANSFORMS wies die Dosis Fingolimod 0,5 mg gegenüber der Dosis Fingolimod 1,25 mg ein besseres Nutzen/Risiko-Profil auf. Aus diesem Grund wurde die 0,5 mg Dosierung zur Zulassung eingereicht. Im Folgenden werden nur Sicherheitsdaten aus der 0,5 mg Dosisgruppe sowie der Vergleichsgruppen berichtet. Sicherheitsdaten zur höheren Dosierung sind den Studienberichten zu entnehmen.

Fingolimod wird von Patienten mit schubförmig verlaufender Multipler Sklerose gut vertragen und die Behandlungsabbruchrate ist gering.

In TRANSFORMS wurden weniger UE für Fingolimod 0,5 mg berichtet als für IFN- β -1a i. m., die zweckmäßige Vergleichstherapie. Während der zweijährigen Studie FREEDOMS wurde die Behandlung aufgrund von UE bei ähnlichen Anteilen von Patienten abgebrochen, unter Fingolimod 0,5 mg bei 7,5 %, unter Placebo bei 7,7 %. Die Inzidenz der UE ist ähnlich in den Behandlungsgruppen mit Fingolimod, IFN- β -1a i. m. und Placebo und die meisten Ereignisse sind von leichter bis mäßiger Intensität. Die Inzidenz der in Studien der Phase III berichteten schwerwiegenden unerwünschten Ereignisse (SUE) war relativ niedrig.

Häufige UE

Häufige UE in beiden Studien (Sicherheitspopulation) waren Kopfschmerzen (RR und 95 %-KI in FREEDOMS 1,08 (0,85; 1,38) und TRANSFORMS 1,14 (0,89; 1,47)) und Nasopharyngitis (RR und 95 %-KI in FREEDOMS 0,98 (0,79; 1,23) und TRANSFORMS 1,00 (0,77; 1,31)), die mit einer ähnlichen Häufigkeit von rund 20-28 % (Fingolimod 0,5 mg, Placebo und IFN- β -1a i. m.) berichtet wurden. Infektionen der oberen Atemwege wurden bei 17,3 % der Patienten in der Studie FREEDOMS berichtet (Fingolimod 0,5 mg und Placebo, RR und 95 %-KI: 0,98 (0,73; 1,32)), jedoch betrug die berichtete Gesamtinzidenz in TRANSFORMS um 7 % (Fingolimod 0,5 mg und IFN- β -1a i. m.; RR und 95 %-KI: 1,15 (0,70; 1,90)). Während der zweijährigen Studie FREEDOMS waren Kopfschmerzen, auffällige Leberfunktionstests, Grippe, Diarrhö, Husten und Rückenschmerzen diejenigen UE (Inzidenz > 10 %), die unter Fingolimod bei einem größeren Anteil Patienten berichtet wurden als unter Placebo (≥ 2 % Unterschied zwischen den Gruppen). Analog dazu wurden in der TRANSFORMS-Studie nur Kopfschmerzen (Inzidenz > 10 %) unter Fingolimod bei einem größeren Anteil Patienten berichtet als unter IFN- β -1a i. m. (≥ 2 % Unterschied zwischen den Gruppen), das 95 %-Konfidenzintervall des relativen Risikos zeigt keine Signifikanz (RR 1,14 (0,89; 1,47)). Auffallend war hier außerdem die hohe Rate von grippeähnlichen Symptomen in der IFN- β -1a i. m. Gruppe, die gegenüber 0,5 mg Fingolimod um das 10-fache erhöht war (36,9 % vs. 3,5 %). Das relative Risiko lag hier bei 0,09 für Fingolimod im Vergleich zu IFN- β -1a i. m. (95 %-KI: 0,06; 0,16). Dieser Unterschied zeigte

sich auch in einer ca. 4-fachen Erhöhung der Pyrexie (17,9 % vs. 4,2 %; RR und 95 %-KI: 0,23 (0,14; 0,39)) sowie einer mehr als 3-fachen Erhöhung der Myalgien (10,2 % vs. 3,3 %; RR und 95 %-KI: 0,32 (0,18; 0,57)).

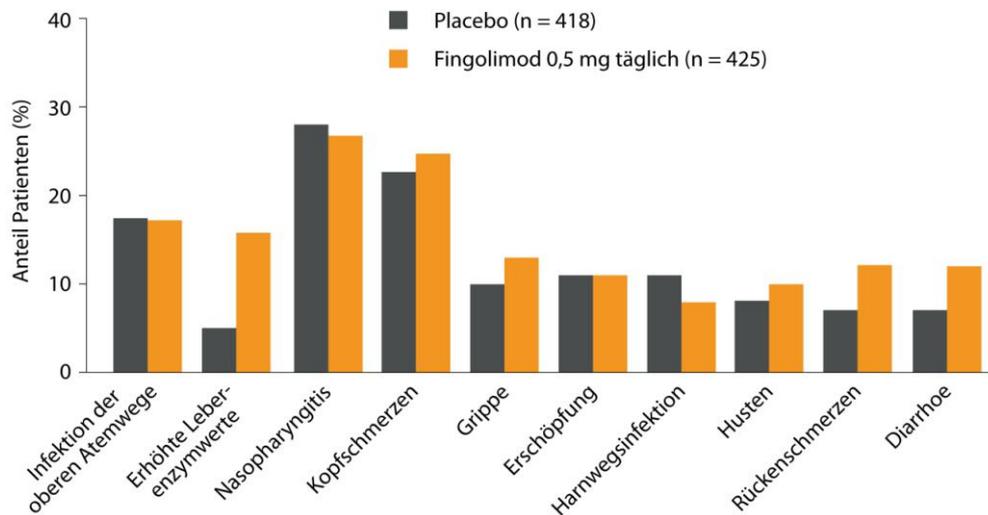


Abbildung 4-19: Unerwünschte Ereignisse bei ≥ 10 % der Patienten nach 24 Monaten in FREEDOMS (Sicherheitspopulation)

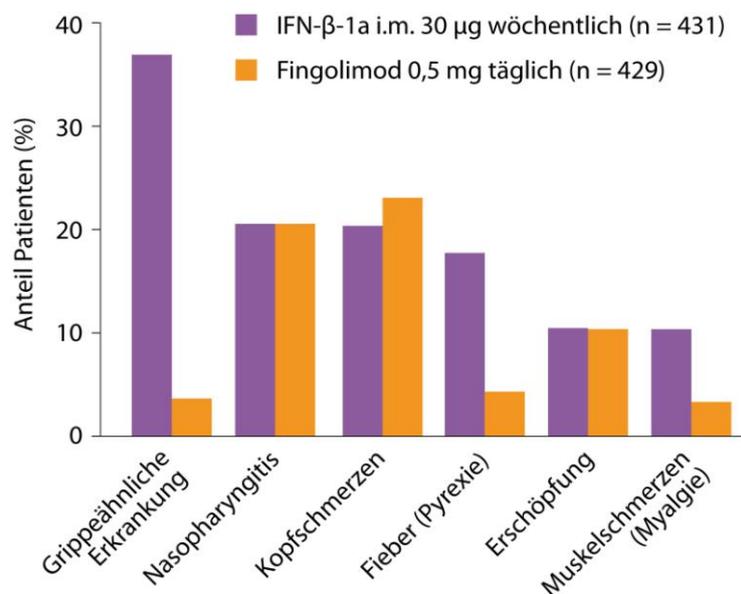


Abbildung 4-20: Unerwünschte Ereignisse bei ≥ 10 % der Patienten nach 12 Monaten in TRANSFORMS (Sicherheitspopulation)

Relevante UE

Bradyarrhythmie/AV-Überleitungsstörungen/ QT-Zeit-Verlängerungen:

In den Studien der Phase III wurde eine vorübergehende, in der Regel asymptomatische Senkung der Herzfrequenz im Mittel um 8 Schläge pro Minute innerhalb einer Stunde nach der ersten Gabe von Fingolimod 0,5 mg beobachtet, die 4-5 Stunden nach Verabreichung das Maximum erreichte. Klinisch relevante Wirkungen auf die Herzfrequenz bildeten sich weitgehend innerhalb von 24 Stunden nach der ersten Gabe von Fingolimod zurück (3). Bei wiederholter Verabreichung wurden keine additiven Effekte festgestellt (4). Als SUE berichtete Bradykardien traten sowohl in FREEDOMS als auch in TRANSFORMS unter Fingolimod 0,5 mg häufiger auf als unter den Vergleichstherapien (4 vs. 1 in FREEDOMS; 2 vs. 0 in TRANSFORMS). Aber weder gegenüber Placebo noch gegenüber IFN- β -1a i. m. konnte ein signifikant erhöhtes Risiko für Bradykardien, AV-Überleitungsstörungen oder QT-Zeit-Verlängerungen abgeleitet werden.

Infektionen:

Angesichts der immunmodulatorischen Eigenschaften von Fingolimod könnte ein erhöhtes Risiko für Infektionen erwartet werden. Jedoch waren die Anteile der Patienten mit Infektionen insgesamt ähnlich für alle Behandlungsgruppen in den Studien TRANSFORMS (51%-52 %) und FREEDOMS (72 %), und die Infektionen waren meist von leichter oder mäßiger Intensität. Infektionen, für die in einer der Behandlungsgruppen aus dem 95%-Konfidenzintervall tatsächlich ein höheres Risiko abzuleiten ist, sind lediglich Bronchitis 2,23 (1,23; 4,03) und Tinea versicolor 4,92 (1,08; 22,31), jeweils gegenüber Placebo. Gegenüber IFN- β -1a i. m. ist kein erhöhtes Risiko für eine bestimmte Infektionsart festzustellen. Es wurden keine opportunistischen Infektionen berichtet.

In der Gruppe mit der höheren Fingolimod-Dosierung 1,25 mg, die nicht Gegenstand der Zulassung ist, wurden zwei Infektionen mit Todesfolge während der Studie TRANSFORMS berichtet: ein Fall primärer disseminierter Windpocken (Varicella zoster) und ein Fall von Herpes-simplex-Enzephalitis. Allerdings lagen in beiden Fällen weitere Einflussfaktoren vor. Im ersten Fall erfolgte zuvor eine Hochdosis-Steroidgabe sowie gegen ärztlichen Rat durch die Patientin selbst ein Dosis-Tapering der Steroide. Die Patientin war dennoch gegen ärztlichen Rat weiterhin in einer Kindertagesstätte, in der Windpockenfälle berichtet wurden, tätig. In der Anamnese war keine frühere Infektion und auch keine Impfung berichtet worden, die Varizella-Zoster-IgG-Titerbestimmung war negativ. Im zweiten Fall wurde nach Auftreten der ersten Symptome der Infektion zunächst irrtümlich mit Antiepileptika und anschließend mit Hochdosis-Steroid behandelt (Behandlung erfolgte nicht im Studienzentrum), bevor 7 Tage nach Auftreten der ersten Symptome am Studienzentrum die korrekte Diagnose der viralen Enzephalitis mittels PCR gestellt und eine entsprechende Behandlung eingeleitet wurde. Die Gesamtexposition gegenüber Fingolimod beträgt jetzt über 10.000 Patientenjahre, daher könnten diese tödlichen Infektionen Einzelfälle darstellen, bei denen ein Zusammenhang mit Fingolimod weder bestätigt noch dementiert werden kann.

Blutdruck:

Geringe Anstiege des mittleren arteriellen Blutdrucks um 2-3 mmHg wurden während der ersten 6 Monate der Behandlung beobachtet. Diese Anstiege blieben für die Dauer der 12-monatigen Studie TRANSFORMS und der 24-monatigen Studie FREEDOMS stabil. Über die Dauer der Studien war aus dem 95 %-Konfidenzintervall weder gegenüber Placebo noch gegenüber IFN- β -1a i. m. ein erhöhtes Risiko für einen Anstieg des Blutdrucks unter Fingolimod ableitbar.

Leberwerterhöhungen:

Sowohl in TRANSFORMS als auch in FREEDOMS wurde eine Erhöhung der Leberwerte festgestellt, dabei kam es bei 8 % der Patienten unter Fingolimod 0,5 mg gegenüber 2 % der Placebo-Patienten zu einem Anstieg der Lebertransaminasen auf das 3-fache der Obergrenze des Normalwerts (ULN) und darüber. Ein Anstieg auf mehr als das 5-fache der ULN trat bei 2 % der Fingolimod- und 1 % der Placebo-Patienten auf. In den Studien wurde Fingolimod abgesetzt, falls der Anstieg das 5-fache des ULN überstieg. Bei erneuter Exposition gegenüber Fingolimod traten bei einigen Patienten erneut erhöhte Lebertransaminasen-Werte auf, was auf einen Zusammenhang mit Fingolimod hindeutet. Die Mehrzahl dieser Erhöhungen trat innerhalb von 3-4 Monaten auf. Die Serumtransaminase-Werte normalisierten sich innerhalb von rund 2 Monaten nach Absetzen von Fingolimod. Das relative Risiko für Leberwerterhöhungen verschiedener Art ist unter Fingolimod gegenüber Placebo beziehungsweise IFN- β -1a i. m. um das 3- bis 4-fache erhöht.

Lymphopenie/ Leukopenie

Ein zentraler pharmakodynamischer Effekt von Fingolimod ist die dosisabhängige Reduktion der peripheren Lymphozytenzahl bis auf 20-30 % des Ausgangswertes. Dies ist auf das reversible Zurückhalten (Sequestrierung) von Lymphozyten im Lymphgewebe zurückzuführen. Aus der FREEDOMS-Studie ist unter Fingolimod ein erhöhtes Risiko für Lymphopenien und Leukopenien im Vergleich zu Placebo ableitbar. In der TRANSFORMS-Studie trat unter Fingolimod nur ein Fall einer Lymphopenie auf. Leukopenien traten unter Fingolimod häufiger auf, ein erhöhtes relatives Risiko gegenüber IFN- β -1a i. m. lässt sich aber nicht belegen.

Makulaödem:

Umfassende Augenuntersuchungen wurden in allen klinischen Studien zu Fingolimod spezifisch durchgeführt, um das mögliche Auftreten von Makulaödem zu monitorieren. Diese traten im gesamten Studienprogramm bei weniger als 1 % der mit Fingolimod behandelten Patienten auf. In FREEDOMS wurde unter Fingolimod 0,5 mg kein Makulaödem berichtet, in TRANSFORMS bei 0,5 % der Patienten. Alle berichteten Fälle von

Makulaödemem waren nach Absetzen von Fingolimod zumeist ohne die Notwendigkeit weiterer Interventionen reversibel.

Maligne Erkrankungen:

In der Studie FREEDOMS traten die Malignitäten der Haut mit ähnlicher Häufigkeit in den Behandlungsgruppen auf, 4 in der Gruppe mit Fingolimod 0,5 mg und 4 in der Placebo-Gruppe. In TRANSFORMS wurden 6 Hautkarzinome in der Gruppe mit Fingolimod 0,5 mg und 2 in der IFN- β -1a i. m.-Gruppe berichtet. Die Hautkarzinome wurden teilweise bereits bei der ersten der regelmäßigen Untersuchungen durch einen Dermatologen diagnostiziert, so dass vorbestehende Anomalien nicht ausgeschlossen werden können. Alle Malignitäten der Haut wurden exzidiert und erfolgreich behandelt.

Die Anzahl der in FREEDOMS und TRANSFORMS beobachteten Brustkrebsfälle entsprach der Inzidenz, die bei einer zu 70 % weiblichen Patientenpopulation im Alter von 20-55 Jahren erwartet wird (38). 5 Fälle wurden in den Studien FREEDOMS (3 unter Placebo) und TRANSFORMS (2 in der Fingolimod 0,5 mg-Gruppe) beobachtet; die 2 Brustkrebsfälle unter Fingolimod wurden innerhalb von 4 Monaten nach Eintritt in die Studie diagnostiziert, was darauf hindeutet, dass sie wahrscheinlich bereits vorher bestanden. Diesen Daten zufolge liegt keine ersichtliche Erhöhung des Brustkrebsrisikos unter Fingolimod-Behandlung vor.

Insgesamt wurden maligne Erkrankungen in den beiden Phase III-Studien selten berichtet; die Fälle traten in TRANSFORMS häufiger unter Fingolimod (12 Patienten; 2,8 %) als unter IFN- β -1a i. m. auf (3 Patienten; 0,7 %), waren jedoch in FREEDOMS unter Fingolimod weniger häufig (5 Patienten; 1,2 %) als unter Placebo (10 Patienten; 2,4 %). Ein erhöhtes relatives Risiko gegenüber IFN- β -1a i. m. ist zwar theoretisch aus den isolierten Studiendaten ableitbar, allerdings zeigen integrierte Sicherheitsdaten (s. Modul 2.7.4 des Common Technical Document) aus FREEDOMS und TRANSFORMS bis zu 12 Monaten, dass bösartige Neubildungen in ähnlichen Anteilen bei den mit Fingolimod behandelten Patienten (0,5 mg: 1,2 %;) und den Patienten, die Vergleichsbehandlungen erhielten, auftraten (IFN- β -1a i. m.: 0,5 %; Placebo: 1,4 %). Ein erhöhtes Risiko wurde von den Zulassungsbehörden nicht gesehen.

Das Sicherheitsprofil in den relevanten Subpopulationen stimmt weitestgehend mit der Gesamtpopulation überein. Es wurden keine bedeutsam von der Gesamtpopulation abweichenden Daten festgestellt. Die Konfidenzintervalle aus den Auswertungen der Subpopulationen sind relativ weit. Ein Vergleich der Daten zwischen Gesamt- und relevanten Subpopulationen ist daher schwer interpretierbar.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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 Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Bei den beiden vorliegenden Phase III-Studien handelt es sich um verschiedene Studientypen (aktiv- beziehungsweise Placebo-kontrolliert), daher erscheint eine Zusammenfassung mittels Meta-Analyse wenig sinnvoll.

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

4.3.1.3.2 Subgruppenanalysen – RCT

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

Beschreiben Sie die Ergebnisse von Subgruppenanalysen (einschließlich der Interaktionsterme). Stellen Sie dabei die Ergebnisse in den Subgruppen zunächst für jede einzelne Studie in tabellarischer Form dar. Diese Anforderung gilt sowohl für Subgruppenanalysen auf Basis individueller Patientendaten als auch für solche auf Basis aggregierter Daten.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse 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.

Im Folgenden werden die Subgruppenanalysen zu Krankheitsschwere und Geschlecht dargestellt. Von der Darstellung weiterer Subgruppenanalysen wird abgesehen, da sie als nicht relevant und nicht aussagekräftig angesehen werden (siehe Abschnitt 4.2.5.5). Für den Endpunkt „Unerwünschte Ereignisse“ werden keine Subgruppenanalysen dargestellt, da hiervon kein zusätzlicher Erkenntnisgewinn erwartet werden kann. Eine Auswertung einzelner relevanter UE-Kategorien ist aufgrund der ohnehin schon geringen Ereigniszahlen nicht weiter aussagekräftig.

Endpunkte „Schübe“Tabelle 4-29: Operationalisierung von „Schübe“^{aa}

Studie	Operationalisierung
FREEDOMS	<p>Aggregierte jährliche Schubrate^c: Gesamtzahl aller bestätigten Schübe aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25</p> <hr/> <p>Anteil schubfreier Patienten zu Monat 24 (Kaplan-Meier Schätzer aus „Zeit bis zum ersten bestätigten Schub“)</p> <hr/> <p>Schubcharakteristika: Anzahl der milden, mittelschweren und schweren Schübe^b; Anzahl der Schübe, die mit Steroiden behandelt wurden; Anzahl der Schübe, die eine Hospitalisierung erforderten</p>
TRANSFORMS	<p>Aggregierte jährliche Schubrate^c: Gesamtzahl aller bestätigten Schübe aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25</p> <hr/> <p>Anteil schubfreier Patienten zu Monat 12 (Kaplan-Meier Schätzer aus „Zeit bis zum ersten bestätigten Schub“)</p> <hr/> <p>Schubcharakteristika: Anzahl der milden, mittelschweren und schweren Schübe^b; Anzahl der Schübe, die mit Steroiden behandelt wurden; Anzahl der Schübe, die eine Hospitalisierung erforderten</p>
<p>a: Ein Schub ist allgemein definiert als neue, verschlechterte oder erneut auftretende neurologische Symptome, die mindestens 30 Tage nach Beginn des vorausgegangenen Schubs auftraten, mindestens 24 Stunden anhielten, ohne Fieber oder Infektion verliefen.</p> <p>Ein bestätigter Schub liegt vor, wenn die allgemeinen Symptome von einem Anstieg um $\geq 0,5$ EDSS-Punkte oder um 1 Punkt in zwei Funktionssystem-Werten oder um 2 Punkte in einem Funktionssystem-Wert (mit Ausnahme von Darm/Blase und psychischen Systemen) begleitet waren.</p> <p>b: Ein milder Schub ist definiert als EDSS-Anstieg um 0,5 Punkte oder 1 Punkt in 1 bis 3 Funktionswerten; ein mittelschwerer Schub ist definiert als EDSS-Anstieg um 1 oder 2 Punkte, eine Änderung um 2 Punkte in 1 oder 2 Funktionswerten oder 1 Punkt Änderung in 4 oder mehr Funktionswerten; ein schwerer Schub ist definiert als ein Schub der die Kriterien eines mittelschweren Schubes übersteigt.</p> <p>c: Aus Gründen der Übersichtlichkeit im Folgenden nur dargestellt als „rate ratio“</p>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-30: Bewertung des Verzerrungspotenzials für „Schübe“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
Aggregierte jährliche Schubrate						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
Anteil schubfreier Patienten						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
Schubcharakteristika: Schubschwere						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
Schubcharakteristika: Schübe, die einen Krankenhausaufenthalt erfordern						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
Schubcharakteristika: Schübe, die eine Steroidbehandlung erfordern						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial für die schubbezogenen Endpunkte (alle Operationalisierungen) für beide Studien als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in beiden Fällen als „niedrig“ eingestuft wurde, jeweils der EDSS-Rater (unabhängiger beurteilender Arzt) verblindet war, das ITT-Prinzip adäquat umgesetzt wurde und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Faktoren gab.

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 4-31: Ergebnisse für „Jährliche Schubrate“ aus RCT mit dem zu bewertenden Arzneimittel (ITT)

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS^{a,b}	Fingolimod 0,5 mg (N=425)	Placebo (N=418)
Subgruppenanalyse Geschlecht		
Männer	n=129	n=120
ARR ratio geg. Placebo (95 %-KI)	0,328 (0,214; 0,503)	
Vergleich zu Placebo	p<0,001	
Frauen	n=296	n=298
ARR ratio geg. Placebo (95 %-KI)	0,500 (0,389; 0,643)	
Vergleich zu Placebo	p<0,001	
Interaktionstest	p=0,212	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0- 3,5	n=363	n=346
ARR ratio geg. Placebo (95 %-KI)	0,475 (0,372; 0,606)	
Vergleich zu Placebo	p<0,001	
EDSS > 3,5	n=62	n=72
ARR ratio geg. Placebo (95 %-KI)	0,348 (0,216; 0,561)	
Vergleich zu Placebo	p<0,001	
Interaktionstest	p=0,414	
TRANSFORMS^{a,b}	Fingolimod 0,5 mg (N=429)	IFN-β-1a i. m. (N=431)
Subgruppenanalyse Geschlecht		
Männer	n=148	n=139
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	0,608 (0,378; 0,978)	
Vergleich zu IFN-β-1a i. m.	p=0,040	
Frauen	n=281	n=292
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	0,443 (0,311; 0,631)	
Vergleich zu IFN-β-1a i. m.	p<0,001	
Interaktionstest	p=0,207	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0 – 3,5	n=362	n=371
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	0,458 (0,332; 0,631)	
Vergleich zu IFN-β-1a i. m.	p<0,001	
EDSS > 3,5	n=67	n=60
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	0,572 (0,314; 1,044)	
Vergleich zu IFN-β-1a i. m.	p=0,069	
Interaktionstest	p=0,898	
a: Effektschätzer und p-Werte für Fingolimod gegenüber den Vergleichsbehandlungen wurden mithilfe des negativen Binomialregressionsmodells berechnet, nach Behandlungsgruppe adjustiert.		
b: Der Interaktionstest (Interaktion Behandlungsgruppe/Geschlecht und Behandlungsgruppe/EDSS) wurde mithilfe des negativen Binomialregressionsmodells berechnet, nach Behandlungsgruppe, Geschlecht und EDSS adjustiert.		

Tabelle 4-32: Ergebnisse für „Jährliche Schubrate“ aus RCT mit dem zu bewertenden Arzneimittel (IFN Non-Responder Def I)

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS^{a,b}	Fingolimod 0,5 mg (N=425)	Placebo (N=418)
Subgruppenanalyse Geschlecht		
Männer	n=19	n=14
ARR ratio geg. Placebo (95 %-KI)	0,320 (0,126; 0,812)	
Vergleich zu Placebo	p=0,016	
Frauen	n=38	n=38
ARR ratio geg. Placebo (95 %-KI)	0,393 (0,187; 0,828)	
Vergleich zu Placebo	p=0,014	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0- 3,5	n=49	n=31
ARR ratio geg. Placebo (95 %-KI)	0,654 (0,323; 1,323)	
Vergleich zu Placebo	p=0,237	
EDSS > 3,5	n=8	n=21
ARR ratio geg. Placebo (95 %-KI)	0,075 (0,009; 0,588)	
Vergleich zu Placebo	p=0,014	
Interaktionstest	nicht bestimmbar ^c	
TRANSFORMS^{a,b}	Fingolimod 0,5 mg (N=429)	IFN-β-1a i. m. (N=431)
Subgruppenanalyse Geschlecht		
Männer	n=44	n=52
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	0,465 (0,221; 0,976)	
Vergleich zu IFN-β-1a i. m.	p=0,043	
Frauen	n=116	n=97
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	0,361 (0,206; 0,632)	
Vergleich zu IFN-β-1a i. m.	p<0,001	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0 – 3,5	n=129	n=125
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	0,326 (0,193; 0,551)	
Vergleich zu IFN-β-1a i. m.	p<0,001	
EDSS > 3,5	n=31	n=24
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	0,615 (0,245; 1,543)	
Vergleich zu IFN-β-1a i. m.	p=0,301	
Interaktionstest	nicht bestimmbar ^c	
a: Effektschätzer und p-Werte für Fingolimod gegenüber den Vergleichsbehandlungen wurden mithilfe des negativen Binomialregressionsmodells berechnet, nach Behandlungsgruppe adjustiert.		
b: Der Interaktionstest (Interaktion Behandlungsgruppe/Geschlecht und Behandlungsgruppe/EDSS) wurde mithilfe des negativen Binomialregressionsmodells berechnet, nach Behandlungsgruppe, Geschlecht und EDSS adjustiert.		

c: Wegen kleiner Patientenzahlen in den Subpopulationen konnten keine komplexen Interaktionsmodelle durchgeführt werden

Tabelle 4-33: Ergebnisse für „Jährliche Schubrate“ aus RCT mit dem zu bewertenden Arzneimittel (IFN Non-Responder Def II)

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS^{a,b}	Fingolimod 0,5 mg (N=425)	Placebo (N=418)
Subgruppenanalyse Geschlecht		
Männer	n=19	n=19
ARR ratio geg. Placebo (95 %-KI)	0,302 (0,093; 0,985)	
Vergleich zu Placebo	p=0,047	
Frauen	n=41	n=35
ARR ratio geg. Placebo (95 %-KI)	0,281 (0,144; 0,549)	
Vergleich zu Placebo	p<0,001	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere		
EDSS 0- 3,5	n=51	n=34
ARR ratio geg. Placebo (95 %-KI)	0,483 (0,226; 1,031)	
Vergleich zu Placebo	p=0,060	
EDSS > 3,5	n=9	n=20
ARR ratio geg. Placebo (95 %-KI)	0,063 (0,009; 0,467)	
Vergleich zu Placebo	p=0,007	
Interaktionstest	nicht bestimmbar ^c	
TRANSFORMS^{a,b}	Fingolimod 0,5 mg (N=429)	IFN-β-1a i. m. (N=431)
Subgruppenanalyse Geschlecht		
Männer	n=44	n=45
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	0,608 (0,292; 1,265)	
Vergleich zu IFN-β-1a i. m.	p=0,183	
Frauen	n=122	n=93
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	0,316 (0,177; 0,565)	
Vergleich zu IFN-β-1a i. m.	p<0,001	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0 – 3,5	n=134	n=115
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	0,312 (0,181; 0,537)	
Vergleich zu IFN-β-1a i. m.	p<0,001	
EDSS > 3,5	n=32	n=23
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	0,671 (0,276; 1,631)	
Vergleich zu IFN-β-1a i. m.	p=0,379	
Interaktionstest	nicht bestimmbar ^c	
a: Effektschätzer und p-Werte für Fingolimod gegenüber den Vergleichsbehandlungen wurden mithilfe des negativen Binomialregressionsmodells berechnet, nach Behandlungsgruppe adjustiert.		
b: Der Interaktionstest (Interaktion Behandlungsgruppe/Geschlecht und Behandlungsgruppe/EDSS) wurde mithilfe des negativen Binomialregressionsmodells berechnet, nach Behandlungsgruppe, Geschlecht und EDSS adjustiert.		

c: Wegen kleiner Patientenzahlen in den Subpopulationen konnten keine komplexen Interaktionsmodelle durchgeführt werden

Tabelle 4-34: Ergebnisse für „Jährliche Schubrate“ aus RCT mit dem zu bewertenden Arzneimittel (Patienten mit schnell fortschreitender schwerer RRMS)

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS^{a,b}	Fingolimod 0,5 mg (N=425)	Placebo (N=418)
Subgruppenanalyse Geschlecht		
Männer	n=15	n=13
ARR ratio geg. Placebo (95 %-KI)	0,424 (0,133; 1,359)	
Vergleich zu Placebo	p=0,149	
Frauen	n=33	n=24
ARR ratio geg. Placebo (95 %-KI)	0,298 (0,151; 0,591)	
Vergleich zu Placebo	p<0,001	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere		
EDSS 0- 3,5	n=41	n=33
ARR ratio geg. Placebo (95 %-KI)	0,328 (0,169; 0,636)	
Vergleich zu Placebo	p<0,001	
EDSS > 3,5	n=7	n=4
ARR ratio geg. Placebo (95 %-KI)	0,337 (0,094; 1,205)	
Vergleich zu Placebo	p=0,094	
Interaktionstest	nicht bestimmbar ^c	
TRANSFORMS^{a,b}	Fingolimod 0,5 mg (N=429)	IFN-β-1a i. m. (N=431)
Subgruppenanalyse Geschlecht		
Männer	n=7	n=7
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	1,087 (0,118; 10,04)	
Vergleich zu IFN-β-1a i. m.	p=0,942	
Frauen	n=20	n=23
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	0,644 (0,156; 2,657)	
Vergleich zu IFN-β-1a i. m.	p=0,543	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0 – 3,5	n=26	n=25
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	0,798 (0,166; 3,839)	
Vergleich zu IFN-β-1a i. m.	p=0,779	
EDSS > 3,5		
ARR ratio geg. IFN-β-1a i. m. (95 %-KI)	nicht bestimmbar ^d	
Vergleich zu IFN-β-1a i. m.	nicht bestimmbar ^d	
Interaktionstest	nicht bestimmbar ^c	
a: Effektschätzer und p-Werte für Fingolimod gegenüber den Vergleichsbehandlungen wurden mithilfe des negativen Binomialregressionsmodells berechnet, nach Behandlungsgruppe adjustiert.		
b: Der Interaktionstest (Interaktion Behandlungsgruppe/Geschlecht und Behandlungsgruppe/EDSS) wurde mithilfe des negativen Binomialregressionsmodells berechnet, nach Behandlungsgruppe, Geschlecht und EDSS adjustiert.		

c: Wegen kleiner Patientenzahlen in den Subpopulationen konnten keine komplexen Interaktionsmodelle durchgeführt werden

d: Wegen kleiner Patientenzahlen in den Subpopulationen konnten keine Schätzer berechnet werden

Tabelle 4-35: Ergebnisse für „Anteil schubfreier Patienten“ aus RCT mit dem zu bewertenden Arzneimittel (ITT)

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS^{a,b}	Fingolimod 0,5 mg (N=425)	Placebo (N=418)
Subgruppenanalyse Geschlecht		
Männer	n=129	n=120
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	74,1 (3,96)	42,9 (4,66)
95 %-KI des KM-Schätzers	(65,38; 80,94)	(33,65; 51,78)
Hazard Ratio (95 %-KI)	0,37 (0,24; 0,57)	
Vergleich gegen Placebo	p<0,001	
Frauen	n=296	n=298
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	68,8 (2,73)	46,7 (2,99)
95 %-KI des KM-Schätzers	(63,10; 73,82)	(40,77; 52,45)
Hazard Ratio (95 %-KI)	0,51 (0,39; 0,66)	
Vergleich gegen Placebo	p<0,001	
Interaktionstest	p=0,339	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0 – 3,5	n=363	n=346
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	70,7 (2,44)	50,0 (2,77)
95 %-KI des KM-Schätzers	(65,59; 75,15)	(44,45; 55,26)
Hazard Ratio (95 %-KI)	0,52 (0,40; 0,66)	
Vergleich gegen Placebo	p<0,001	
EDSS > 3,5	n=62	n=72
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	68,7 (5,96)	24,0 (5,32)
95 %-KI des KM-Schätzers	(55,43; 78,79)	(14,45; 34,90)
Hazard Ratio (95 %-KI)	0,29 (0,17; 0,49)	
Vergleich gegen Placebo	p<0,001	
Interaktionstest	p=0,173	
TRANSFORMS^{a,b}		
	Fingolimod 0,5 mg (N=429)	IFN-β-1a i. m. (N=431)
Subgruppenanalyse Geschlecht		
Männer	n=148	n=139
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	82,0 (3,20)	71,2 (3,95)
95 %-KI des KM-Schätzers	(74,69; 87,38)	(62,63; 78,16)
Hazard Ratio (95 %-KI)	0,61 (0,37; 1,00)	
Vergleich gegen IFN-β-1a i. m.	p=0,049	
Frauen	n=281	n=292
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	83,0 (2,26)	68,4 (2,78)

95 %-KI des KM-Schätzers	(78,01; 86,94)	(62,60; 73,50)
Hazard Ratio (95 %-KI)	0,50 (0,35; 0,70)	
Vergleich gegen IFN- β -1a i. m.	p<0,001	
Interaktionstest	p=0,166	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0 – 3,5	n=362	n=371
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	84,5 (1,93)	70,8 (2,41)
95 %-KI des KM-Schätzers	(80,24; 87,85)	(65,75; 75,19)
Hazard Ratio (95 %-KI)	0,50 (0,37; 0,70)	
Vergleich gegen IFN- β -1a i. m.	p<0,001	
EDSS > 3,5	n=67	n=60
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	73,1 (5,43)	59,4 (6,74)
95 %-KI des KM-Schätzers	(60,72; 82,10)	(44,98; 71,17)
Hazard Ratio (95 %-KI)	0,58 (0,31; 1,07)	
Vergleich gegen IFN- β -1a i. m.	p=0,082	
Interaktionstest	p=0,943	
<p>a: HRs und <i>p</i>-Werte gelten für Fingolimod vs. Placebo oder IFN-β-1a i. m.; die <i>p</i>-Werte für die HRs wurden mittels des proportionalen Risikomodells nach Cox berechnet, adjustiert nach Behandlung.</p> <p>b: Der Interaktionstest (Interaktion Behandlungsgruppe/Geschlecht und Behandlungsgruppe/EDSS) wurde mithilfe des proportionalen Risikomodells nach Cox berechnet, nach Behandlungsgruppe, Geschlecht und EDSS adjustiert.</p>		

Tabelle 4-36: Ergebnisse für „Anteil schubfreier Patienten“ aus RCT mit dem zu bewertenden Arzneimittel (IFN Non-Responder Def I)

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS^{a,b}	Fingolimod 0,5 mg (N=425)	Placebo (N=418)
Subgruppenanalyse Geschlecht		
Männer	n=19	n=14
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	71,4 (10,84)	30,6 (13,48)
95 %-KI des KM-Schätzers	(44,30; 87,02)	(8,59; 56,50)
Hazard Ratio (95 %-KI)	0,35 (0,12; 1,05)	
Vergleich gegen Placebo	p=0,061	
Frauen	n=38	n=38
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	66,7 (7,86)	40,8 (8,40)
95 %-KI des KM-Schätzers	(48,86; 79,56)	(24,51; 56,48)
Hazard Ratio (95 %-KI)	0,43 (0,21; 0,87)	
Vergleich gegen Placebo	p=0,020	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0 – 3,5	n=49	n=31
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	65,4 (7,01)	51,7 (9,46)
95 %-KI des KM-Schätzers	(49,85; 77,20)	(32,10; 68,18)
Hazard Ratio (95 %-KI)	0,63 (0,31; 1,29)	
Vergleich gegen Placebo	p=0,206	
EDSS > 3,5	n=8	n=21
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	85,7 (13,23)	17,4 (8,94)
95 %-KI des KM-Schätzers	(33,41; 97,86)	(4,47; 37,43)
Hazard Ratio (95 %-KI)	0,09 (0,01; 0,71)	
Vergleich gegen Placebo	p=0,022	
Interaktionstest	nicht bestimmbar ^c	
TRANSFORMS^{a,b}	Fingolimod 0,5 mg (N=429)	IFN-β-1a i. m. (N=431)
Subgruppenanalyse Geschlecht		
Männer	n=44	n=52
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	81,7 (5,85)	57,1 (7,32)
95 %-KI des KM-Schätzers	(66,70; 90,40)	(41,56; 69,89)
Hazard Ratio (95 %-KI)	0,40 (0,18; 0,92)	
Vergleich gegen IFN-β-1a i. m.	p=0,031	
Frauen	n=116	n=97
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	84,3 (3,40)	63,8 (4,98)

95 %-KI des KM-Schätzers	(76,22; 89,79)	(53,15; 72,63)
Hazard Ratio (95 %-KI)	0,38 (0,22; 0,66)	
Vergleich gegen IFN- β -1a i. m.	p<0,001	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0 – 3,5	n=129	n=125
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	85,9 (3,08)	62,6 (4,43)
95 %-KI des KM-Schätzers	(78,53; 90,86)	(53,24; 70,57)
Hazard Ratio (95 %-KI)	0,34 (0,20; 0,58)	
Vergleich gegen IFN- β -1a i. m.	p<0,001	
EDSS > 3,5	n=31	n=24
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	74,1 (7,90)	55,3 (11,25)
95 %-KI des KM-Schätzers	(54,74; 86,09)	(31,34; 73,96)
Hazard Ratio (95 %-KI)	0,49 (0,19; 1,28)	
Vergleich gegen IFN- β -1a i. m.	p=0,146	
Interaktionstest	nicht bestimmbar ^c	
<p>a: HRs und <i>p</i>-Werte gelten für Fingolimod vs. Placebo oder IFN-β-1a i. m.; die <i>p</i>-Werte für die HRs wurden mittels des proportionalen Risikomodells nach Cox berechnet, adjustiert nach Behandlung.</p> <p>b: Der Interaktionstest (Interaktion Behandlungsgruppe/Geschlecht und Behandlungsgruppe/EDSS) wurde mithilfe des proportionalen Risikomodells nach Cox berechnet, nach Behandlungsgruppe, Geschlecht und EDSS adjustiert.</p> <p>c: Wegen kleiner Patientenzahlen in den Subpopulationen konnten keine komplexen Interaktionsmodelle durchgeführt werden</p>		

Tabelle 4-37: Ergebnisse für „Anteil schubfreier Patienten, “ aus RCT mit dem zu bewertenden Arzneimittel (IFN Non-Responder Def II)

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS^{a,b}	Fingolimod 0,5 mg (N=425)	Placebo (N=418)
Subgruppenanalyse Geschlecht		
Männer	n=19	n=19
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	77,0 (10,10)	42,6 (12,32)
95 %-KI des KM-Schätzers	(49,72; 90,72)	(19,08; 64,48)
Hazard Ratio (95 %-KI)	0,34 (0,11; 1,09)	
Vergleich gegen Placebo	p=0,071	
Frauen	n=41	n=35
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	74,5 (6,97)	27,9 (8,27)
95 %-KI des KM-Schätzers	(57,79; 85,42)	(13,38; 44,51)
Hazard Ratio (95 %-KI)	0,25 (0,12; 0,53)	
Vergleich gegen Placebo	p<0,001	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0 – 3,5	n=51	n=34
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	73,1 (6,38)	44,2 (9,42)
95 %-KI des KM-Schätzers	(58,24; 83,43)	(25,65; 61,32)
Hazard Ratio (95 %-KI)	0,43 (0,21; 0,88)	
Vergleich gegen Placebo	p=0,021	
EDSS > 3,5	n=9	n=20
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	87,5 (11,69)	12,5 (8,11)
95 %-KI des KM-Schätzers	(38,70; 98,14)	(2,16; 32,38)
Hazard Ratio (95 %-KI)	0,07 (0,01; 0,54)	
Vergleich gegen Placebo	p=0,011	
Interaktionstest	nicht bestimmbar ^c	
TRANSFORMS^{a,b}	Fingolimod 0,5 mg (N=429)	IFN-β-1a i. m. (N=431)
Subgruppenanalyse Geschlecht		
Männer	n=44	n=45
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	77,3 (6,32)	63,1 (7,63)
95 %-KI des KM-Schätzers	(61,88; 87,07)	(46,30; 75,96)
Hazard Ratio (95 %-KI)	0,63 (0,28; 1,41)	
Vergleich gegen IFN-β-1a i. m.	p=0,260	
Frauen	n=122	n=93
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	85,8 (3,18)	64,6 (5,05)

95 %-KI des KM-Schätzers	(78,21; 90,95)	(53,73; 73,48)
Hazard Ratio (95 %-KI)	0,35 (0,20; 0,63)	
Vergleich gegen IFN- β -1a i. m.	p<0,001	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0 – 3,5	n=134	n=115
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	86,4 (2,98)	65,2 (4,52)
95 %-KI des KM-Schätzers	(79,29; 91,21)	(55,55; 73,23)
Hazard Ratio (95 %-KI)	0,36 (0,21; 0,62)	
Vergleich gegen IFN- β -1a i. m.	p<0,001	
EDSS > 3,5	n=32	n=23
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	71,9 (7,95)	58,2 (11,46)
95 %-KI des KM-Schätzers	(52,91; 84,26)	(33,12; 76,67)
Hazard Ratio (95 %-KI)	0,63 (0,24; 1,63)	
Vergleich gegen IFN- β -1a i. m.	p=0,341	
Interaktionstest	nicht bestimmbar ^c	
<p>a: HRs und <i>p</i>-Werte gelten für Fingolimod vs. Placebo oder IFN-β-1a i. m.; die <i>p</i>-Werte für die HRs wurden mittels des proportionalen Risikomodells nach Cox berechnet, adjustiert nach Behandlung.</p> <p>b: Der Interaktionstest (Interaktion Behandlungsgruppe/Geschlecht und Behandlungsgruppe/EDSS) wurde mithilfe des proportionalen Risikomodells nach Cox berechnet, nach Behandlungsgruppe, Geschlecht und EDSS adjustiert.</p> <p>c: Wegen kleiner Patientenzahlen in den Subpopulationen konnten keine komplexen Interaktionsmodelle durchgeführt werden.</p>		

Tabelle 4-38: Ergebnisse für „Anteil schubfreier Patienten“ aus RCT mit dem zu bewertenden Arzneimittel (Patienten mit schnell fortschreitender schwerer RRMS)

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS^{a,b}	Fingolimod 0,5 mg (N=425)	Placebo (N=418)
Subgruppenanalyse Geschlecht		
Männer	n=15	n=13
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	66,7 (12,17)	30,8 (12,80)
95 %-KI des KM-Schätzers	(37,53; 84,56)	(9,50; 55,43)
Hazard Ratio (95 %-KI)	0,47 (0,16; 1,42)	
Vergleich gegen Placebo	p=0,181	
Frauen	n=33	n=24
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	71,8 (7,99)	23,7 (9,16)
95 %-KI des KM-Schätzers	(52,75; 84,25)	(8,74; 42,72)
Hazard Ratio (95 %-KI)	0,26 (0,12; 0,60)	
Vergleich gegen Placebo	p=0,001	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0 – 3,5	n=41	n=33
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	75,0 (6,86)	29,8 (8,25)
95 %-KI des KM-Schätzers	(58,52; 85,73)	(15,01; 46,13)
Hazard Ratio (95 %-KI)	0,28 (0,13; 0,60)	
Vergleich gegen Placebo	p=0,001	
EDSS > 3,5	n=7	n=4
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 24	42,9 (18,70)	0
95 %-KI des KM-Schätzers	(9,78; 73,44)	nicht bestimmbar
Hazard Ratio (95 %-KI)	0,54 (0,13; 2,18)	
Vergleich gegen Placebo	p=0,388	
Interaktionstest	nicht bestimmbar ^c	
TRANSFORMS^{a,b}		
	Fingolimod 0,5 mg (N=429)	IFN-β-1a i. m. (N=431)
Subgruppenanalyse Geschlecht		
Männer	n=7	n=7
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	66,7 (19,25)	85,7 (13,23)
95 %-KI des KM-Schätzers	(19,46; 90,44)	(33,41; 97,86)
Hazard Ratio (95 %-KI)	2,13 (0,19; 23,5)	
Vergleich gegen IFN-β-1a i. m.	p=0,538	
Frauen	n=20	n=23
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	84,7 (8,14)	73,2 (9,37)

95 %-KI des KM-Schätzers	(59,68; 94,80)	(49,80; 87,00)
Hazard Ratio (95 %-KI)	0,54 (0,13; 2,15)	
Vergleich gegen IFN- β -1a i. m.	p=0,381	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0 – 3,5	n=26	n=25
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	83,2 (7,68)	83,8 (7,42)
95 %-KI des KM-Schätzers	(61,14; 93,37)	(62,40; 93,60)
Hazard Ratio (95 %-KI)	0,94 (0,24; 3,76)	
Vergleich gegen IFN- β -1a i. m.	p=0,932	
EDSS > 3,5		
KM-Schätzer für Anteil schubfreier Patienten in % (SE in %) zu Monat 12	nicht bestimmbar ^d	nicht bestimmbar ^d
95 %-KI des KM-Schätzers	nicht bestimmbar ^d	nicht bestimmbar ^d
Hazard Ratio (95 %-KI)	nicht bestimmbar ^d	
Vergleich gegen IFN- β -1a i. m.	nicht bestimmbar ^d	
Interaktionstest	nicht bestimmbar ^c	
a: HRs und <i>p</i> -Werte gelten für Fingolimod vs. Placebo oder IFN- β -1a i. m.; die <i>p</i> -Werte für die HRs wurden mittels des proportionalen Risikomodells nach Cox berechnet, adjustiert nach Behandlung.		
b: Der Interaktionstest (Interaktion Behandlungsgruppe/Geschlecht und Behandlungsgruppe/EDSS) wurde mithilfe des proportionalen Risikomodells nach Cox berechnet, nach Behandlungsgruppe, Geschlecht und EDSS adjustiert.		
c: Wegen kleiner Patientenzahlen in den Subpopulationen konnten keine komplexen Interaktionsmodelle durchgeführt werden.		
d: Wegen kleiner Patientenzahlen in den Subpopulationen konnten keine Schätzer berechnet werden.		

Die Interaktionstests und Subgruppenauswertungen in der ITT-Population, die für beide Studien durchgeführt wurden, geben keinen Hinweis auf einen Einfluss des Geschlechts oder der Krankheitsschwere (definiert über den Baseline EDSS) auf die Wirksamkeit von Fingolimod hinsichtlich der Schubratenreduktion und des Anteils schubfreier Patienten.

In den Subpopulationen sind die Subgruppenanalysen und Interaktionstests mit Vorsicht zu interpretieren, da es sich häufig um sehr kleine Patientenzahlen handelt. Es ist daher kein eindeutiger Hinweis auf einen Einfluss des Geschlechts oder der Krankheitsschwere (definiert über den Baseline EDSS) auf die Wirksamkeit von Fingolimod hinsichtlich der Schubratenreduktion und des Anteils schubfreier Patienten festzustellen.

HTA Tabelle 15.1.1 bis HTA Tabelle 15.4.8 im Anhang 4-H enthalten die deskriptive Darstellung der Subgruppenanalysen für die Schubcharakteristika in den Gesamt- bzw. Subpopulationen. HTA Tabelle 16.1.1 bis HTA Tabelle 16.4.7 in Anhang 4-H enthalten die statistischen Tests (*p*-Werte aus Fishers Exakt-Test; für die Subpopulation der Patienten mit schnell fortschreitender schwerer RRMS in der Subgruppe EDSS > 3,5 konnte wegen zu kleiner Patientenzahlen kein statistischer Test durchgeführt werden) der Subgruppenanalysen für die Gesamt- bzw. Subpopulationen. Die Tabellen sind aus Gründen der Übersichtlichkeit

im Anhang 4-H zu finden und nicht im Hauptteil dargestellt. Ein Modell wurde nicht für angebracht erachtet, da die Daten rein beschreibenden Charakter haben, ein Interaktionstests ist daher nicht zutreffend. Die Subgruppenanalysen zu den Schubcharakteristika in der Gesamtpopulation sowie den Subpopulationen zeigen keine auffälligen Unterschiede zwischen den einzelnen Subgruppen.

Bei den beiden vorliegenden Phase III-Studien handelt es sich um verschiedene Studientypen (aktiv- beziehungsweise Placebo-kontrolliert), daher erscheint eine Zusammenfassung mittels Meta-Analyse wenig sinnvoll.

Endpunkte „Behinderungsprogression“

Tabelle 4-39: Operationalisierung von „Behinderungsprogression“^{ca}

Studie	Operationalisierung
FREEDOMS	Anteil Patienten ohne nach 3 Monaten bestätigter Behinderungsprogression ^a (Kaplan-Meier-Schätzer) zu Monat 24 Mittlere Änderung des EDSS zu Monat 24 gegenüber Baseline Mittlere Änderung des MSFC zu Monat 24 gegenüber Baseline
TRANSFORMS	Anteil Patienten ohne nach 3 Monaten bestätigter Behinderungsprogression ^a (Kaplan-Meier-Schätzer) zu Monat 12 Mittlere Änderung des EDSS zu Monat 12 gegenüber Baseline Mittlere Änderung des MSFC zu Monat 12 gegenüber Baseline
a: Die Behinderungsprogression war definiert als eine Erhöhung des EDSS-Werts um 1 Punkt (um 0,5 Punkte bei einem Baseline-EDSS-Wert $\geq 5,5$), die 3 Monate später ohne Vorliegen eines Schubs bestätigt wurde.	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-40: Bewertung des Verzerrungspotenzials für „Behinderungsprogression“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
Anteil behinderungsprogressionsfreier Patienten						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
Mittlere Änderung des EDSS						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
Mittlere Änderung des MSFC						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial für die Endpunkte zur Behinderung (alle Operationalisierungen) für beide Studien als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in beiden Fällen als „niedrig“ eingestuft wurde, jeweils der EDSS-Rater (unabhängiger beurteilender Arzt) und MSFC-Rater (unabhängiger beurteilender Arzt oder qualifizierte Person) verblindet war, das ITT-Prinzip adäquat umgesetzt wurde und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte gab.

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 4-41: Ergebnisse für „Anteil Patienten ohne Behinderungsprogression“ aus RCT mit dem zu bewertenden Arzneimittel (ITT)

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS^{a,b}	Fingolimod 0,5 mg (N=425)	Placebo (N=418)
Subgruppenanalyse Geschlecht		
Männer	n=129	n=120
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	86,7 (3,10)	75,0 (4,12)
95 %-KI des KM-Schätzers	(79,24; 91,66)	(65,84; 82,06)
Hazard Ratio (95 %-KI)	0,50 (0,27; 0,93)	
Vergleich zu Placebo	p=0,029	
Frauen	n=296	n=298
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	80,6 (2,33)	76,3 (2,55)
95 %-KI des KM-Schätzers	(75,50; 84,69)	(70,89; 80,91)
Hazard Ratio (95 %-KI)	0,80 (0,56; 1,14)	
Vergleich zu Placebo	p=0,217	
Interaktionstest	p=0,411	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0 - 3,5	n=363	n=346
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	82,0 (2,06)	78,4 (2,29)
95 %-KI des KM-Schätzers	(77,52; 85,63)	(73,51; 82,52)
Hazard Ratio (95 %-KI)	0,83 (0,59; 1,16)	
Vergleich zu Placebo	p=0,269	
EDSS > 3,5	n=62	n=72
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	84,3 (4,81)	63,7 (5,95)
95 %-KI des KM-Schätzers	(71,99; 91,54)	(50,82; 74,04)
Hazard Ratio (95 %-KI)	0,35 (0,16; 0,76)	
Vergleich zu Placebo	p=0,008	
Interaktionstest	p=0,022	
TRANSFORMS^{a,b}	Fingolimod 0,5 mg (N=429)	IFN- β -1a i. m. (N=431)
Subgruppenanalyse Geschlecht		
Männer	n=148	n=139
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	93,8 (2,01)	94,0 (2,05)
95 %-KI des KM-Schätzers	(88,38; 96,72)	(88,41; 96,97)

Hazard Ratio (95 %-KI)	1,04 (0,40; 2,70)	
Vergleich zu IFN- β -1a i. m.	p=0,934	
Frauen	n=281	n=292
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	94,2 (1,40)	91,1 (1,70)
95 %-KI des KM-Schätzers	(90,76; 96,43)	(87,13; 93,91)
Hazard Ratio (95 %-KI)	0,64 (0,34; 1,19)	
Vergleich zu IFN- β -1a i. m.	p=0,160	
Interaktionstest	p=0,694	
Subgruppenanalyse Krankheitsschwere (EDSS zu Baseline), (ITT)^b		
EDSS 0,0 – 3,5	n=362	n=371
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	93,8 (1,28)	91,7 (1,46)
95 %-KI des KM-Schätzers	(90,74; 95,88)	(88,29; 94,10)
Hazard Ratio (95 %-KI)	0,73 (0,42; 1,27)	
Vergleich zu IFN- β -1a i. m.	p=0,268	
EDSS > 3,5	n=67	n=60
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	95,5 (2,53)	94,5 (3,09)
95 %-KI des KM-Schätzers	(86,76; 98,53)	(83,90; 98,19)
Hazard Ratio (95 %-KI)	0,82 (0,17; 4,08)	
Vergleich zu IFN- β -1a i. m.	p=0,812	
Interaktionstest	p=0,936	
a: HRs und <i>p</i> -Werte gelten für Fingolimod vs. Placebo oder IFN- β -1a i. m.; die <i>p</i> -Werte für die HRs wurden mittels des proportionalen Risikomodells nach Cox berechnet, adjustiert nach Behandlung.		
b: Der Interaktionstest (Interaktion Behandlungsgruppe/Geschlecht und Behandlungsgruppe/EDSS) wurde mithilfe des proportionalen Risikomodells nach Cox berechnet, nach Behandlungsgruppe, Geschlecht und EDSS adjustiert.		

Tabelle 4-42: Ergebnisse für „Anteil Patienten ohne Behinderungsprogression“ aus RCT mit dem zu bewertenden Arzneimittel (IFN Non-Responder Def I)

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS^{a,b}	Fingolimod 0,5 mg (N=425)	Placebo (N=418)
Subgruppenanalyse Geschlecht		
Männer	n=19	n=14
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	94,4 (5,40)	63,5 (13,08)
95 %-KI des KM-Schätzers	(66,64; 99,20)	(33,12; 82,97)
Hazard Ratio (95 %-KI)	0,14 (0,02; 1,22)	
Vergleich zu Placebo	p=0,075	
Frauen	n=38	n=38
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	78,4 (6,79)	79,7 (6,89)
95 %-KI des KM-Schätzers	(61,32; 88,56)	(61,93; 89,78)
Hazard Ratio (95 %-KI)	1,15 (0,42; 3,18)	
Vergleich zu Placebo	p=0,783	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0- 3,5	n=49	n=31
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	82,9 (5,52)	78,4 (7,82)
95 %-KI des KM-Schätzers	(68,62; 91,05)	(58,12; 89,71)
Hazard Ratio (95 %-KI)	0,90 (0,31; 2,59)	
Vergleich zu Placebo	p=0,842	
EDSS > 3,5	n=8	n=21
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	87,5 (11,69)	69,6 (10,37)
95 %-KI des KM-Schätzers	(38,70; 98,14)	(44,47; 85,09)
Hazard Ratio (95 %-KI)	0,39 (0,05; 3,28)	
Vergleich zu Placebo	p=0,389	
Interaktionstest	nicht bestimmbar ^c	
TRANSFORMS^{a,b}	Fingolimod 0,5 mg (n=429)	IFN-β-1a i. m. (N=431)
Subgruppenanalyse Geschlecht		
Männer	n=44	n=52
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	95,4 (3,18)	96,0 (2,77)
95 %-KI des KM-Schätzers	(82,83; 98,83)	(84,94; 98,98)

Hazard Ratio (95 %-KI)	1,12 (0,16; 7,92)	
Vergleich zu IFN- β -1a i. m.	p=0,913	
Frauen	n=116	n=97
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	93,9 (2,24)	87,9 (3,42)
95 %-KI des KM-Schätzers	(87,62; 97,04)	(79,25; 93,13)
Hazard Ratio (95 %-KI)	0,49 (0,19; 1,28)	
Vergleich zu IFN- β -1a i. m.	p=0,146	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere (EDSS zu Baseline), (ITT)^b		
EDSS 0,0 – 3,5	n=129	n=125
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	93,7 (2,16)	90,8 (2,66)
95 %-KI des KM-Schätzers	(87,79; 96,80)	(83,92; 94,77)
Hazard Ratio (95 %-KI)	0,67 (0,27; 1,67)	
Vergleich zu IFN- β -1a i. m.	p=0,394	
EDSS > 3,5	n=31	n=24
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	96,8 (3,17)	89,9 (6,76)
95 %-KI des KM-Schätzers	(79,23; 99,54)	(65,34; 97,40)
Hazard Ratio (95 %-KI)	0,31 (0,03; 3,47)	
Vergleich zu IFN- β -1a i. m.	p=0,345	
Interaktionstest	nicht bestimmbar ^c	
a: HRs und <i>p</i> -Werte gelten für Fingolimod vs. Placebo oder IFN- β -1a i. m.; die <i>p</i> -Werte für die HRs wurden mittels des proportionalen Risikomodells nach Cox berechnet, adjustiert nach Behandlung.		
b: Der Interaktionstest (Interaktion Behandlungsgruppe/Geschlecht und Behandlungsgruppe/EDSS) wurde mithilfe des proportionalen Risikomodells nach Cox berechnet, nach Behandlungsgruppe, Geschlecht und EDSS adjustiert.		
c: Wegen kleiner Patientenzahlen in den Subpopulationen konnten keine komplexen Interaktionsmodelle durchgeführt werden.		

Tabelle 4-43: Ergebnisse für „Anteil Patienten ohne Behinderungsprogression“ aus RCT mit dem zu bewertenden Arzneimittel (IFN Non-Responder Def II)

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS^{a,b}	Fingolimod 0,5 mg (N=425)	Placebo (N=418)
Subgruppenanalyse Geschlecht		
Männer	n=19	n=19
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	94,4 (5,40)	64,6 (11,76)
95 %-KI des KM-Schätzers	(66,64; 99,20)	(37,22; 82,45)
Hazard Ratio (95 %-KI)	0,15 (0,02; 1,27)	
Vergleich zu Placebo	p=0,082	
Frauen	n=41	n=35
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	80,0 (6,34)	80,4 (7,23)
95 %-KI des KM-Schätzers	(63,92; 89,47)	(61,30; 90,72)
Hazard Ratio (95 %-KI)	1,08 (0,38; 3,12)	
Vergleich zu Placebo	p=0,884	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0- 3,5	n=51	n=34
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	83,6 (5,32)	77,6 (8,12)
95 %-KI des KM-Schätzers	(69,80; 91,44)	(56,62; 89,33)
Hazard Ratio (95 %-KI)	0,86 (0,30; 2,50)	
Vergleich zu Placebo	p=0,788	
EDSS > 3,5	n=9	n=20
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	88,9 (10,48)	68,0 (10,80)
95 %-KI des KM-Schätzers	(43,30; 98,36)	(42,14; 84,21)
Hazard Ratio (95 %-KI)	0,33 (0,04; 2,70)	
Vergleich zu Placebo	p=0,298	
Interaktionstest	nicht bestimmbar ^c	
TRANSFORMS^{a,b}	Fingolimod 0,5 mg (N=429)	IFN-β-1a i. m. (N=431)
Subgruppenanalyse Geschlecht		
Männer	n=44	n=45
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	93,2 (3,80)	97,7 (2,25)
95 %-KI des KM-Schätzers	(80,33; 97,75)	(84,94; 99,68)

Hazard Ratio (95 %-KI)	3,02 (0,31; 29,0)	
Vergleich zu IFN- β -1a i. m.	p=0,339	
Frauen	n=122	n=93
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	95,8 (1,82)	88,6 (3,39)
95 %-KI des KM-Schätzers	(90,28; 98,24)	(79,85; 93,71)
Hazard Ratio (95 %-KI)	0,36 (0,12; 1,06)	
Vergleich zu IFN- β -1a i. m.	p=0,064	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere (EDSS zu Baseline), (ITT)^b		
EDSS 0,0 – 3,5	n=134	n=115
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	94,7 (1,95)	91,8 (2,61)
95 %-KI des KM-Schätzers	(89,20; 97,44)	(84,89; 95,67)
Hazard Ratio (95 %-KI)	0,66 (0,24; 1,76)	
Vergleich zu IFN- β -1a i. m.	p=0,404	
EDSS > 3,5	n=32	n=23
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	96,9 (3,08)	89,4 (7,10)
95 %-KI des KM-Schätzers	(79,82; 99,55)	(63,78; 97,25)
Hazard Ratio (95 %-KI)	0,29 (0,03; 3,19)	
Vergleich zu IFN- β -1a i. m.	p=0,311	
Interaktionstest	nicht bestimmbar ^c	
a: HRs und <i>p</i> -Werte gelten für Fingolimod vs. Placebo oder IFN- β -1a i. m.; die <i>p</i> -Werte für die HRs wurden mittels des proportionalen Risikomodells nach Cox berechnet, adjustiert nach Behandlung.		
b: Der Interaktionstest (Interaktion Behandlungsgruppe/Geschlecht und Behandlungsgruppe/EDSS) wurde mithilfe des proportionalen Risikomodells nach Cox berechnet, nach Behandlungsgruppe, Geschlecht und EDSS adjustiert.		
c: Wegen kleiner Patientenzahlen in den Subpopulationen konnten keine komplexen Interaktionsmodelle durchgeführt werden.		

Tabelle 4-44: Ergebnisse für „Anteil Patienten ohne Behinderungsprogression“ aus RCT mit dem zu bewertenden Arzneimittel (Patienten mit schnell fortschreitender schwerer RRMS)

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS^{a,b}	Fingolimod 0,5 mg (N=425)	Placebo (N=418)
Subgruppenanalyse Geschlecht		
Männer	n=15	n=13
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	85,6 (9,50)	90,9 (8,67)
95 %-KI des KM-Schätzers	(53,27; 96,21)	(50,81; 98,67)
Hazard Ratio (95 %-KI)	1,85 (0,17; 20,5)	
Vergleich zu Placebo	p=0,615	
Frauen	n=33	n=24
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	84,4 (6,42)	72,1 (9,72)
95 %-KI des KM-Schätzers	(66,46; 93,18)	(48,01; 86,41)
Hazard Ratio (95 %-KI)	0,53 (0,16; 1,73)	
Vergleich zu Placebo	p=0,292	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere		
EDSS 0,0- 3,5	n=41	n=33
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	84,9 (5,67)	86,5 (6,29)
95 %-KI des KM-Schätzers	(69,48; 92,93)	(67,88; 94,75)
Hazard Ratio (95 %-KI)	1,13 (0,32; 4,01)	
Vergleich zu Placebo	p=0,849	
EDSS > 3,5	n=7	n=4
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	83,3 (15,21)	25,0 (21,65)
95 %-KI des KM-Schätzers	(27,31; 97,47)	(0,89; 66,53)
Hazard Ratio (95 %-KI)	0,19 (0,02; 1,80)	
Vergleich zu Placebo	p=0,147	
Interaktionstest	nicht bestimmbar ^c	
TRANSFORMS^{a,b}	Fingolimod 0,5 mg (N=429)	IFN-β-1a i. m. (N=431)
Subgruppenanalyse Geschlecht		
Männer	n=7	n=7
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	100,0 ^c	100,0 ^c
95 %-KI des KM-Schätzers	nicht bestimmbar ^e	nicht bestimmbar ^e

Hazard Ratio (95 %-KI) Vergleich zu IFN- β -1a i. m.	nicht bestimmbar ^e nicht bestimmbar ^e	
Frauen	n=20	n=23
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	100,0 ^c	82,4 (8,01)
95 %-KI des KM-Schätzers	nicht bestimmbar ^e	(59,58; 93,0)
Hazard Ratio (95 %-KI) Vergleich zu IFN- β -1a i. m.	nicht bestimmbar ^e nicht bestimmbar ^e	
Interaktionstest	nicht bestimmbar ^c	
Subgruppenanalyse Krankheitsschwere (EDSS zu Baseline), (ITT)^b		
EDSS 0,0 – 3,5	n=26	n=25
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	100,0 ^c	83,8 (7,42)
95 %-KI des KM-Schätzers	nicht bestimmbar ^e	(62,40; 93,60)
Hazard Ratio (95 %-KI) Vergleich zu IFN- β -1a i. m.	nicht bestimmbar ^e nicht bestimmbar ^c	
EDSS > 3,5	n=1	n=5
KM-Schätzer für Anteil Patienten ohne Behinderungsprogression % (SE)	100,0 ^c	100,0 ^c
95 %-KI des KM-Schätzers	nicht bestimmbar ^e	nicht bestimmbar ^e
Hazard Ratio (95 %-KI) Vergleich zu P IFN- β -1a i. m.	nicht bestimmbar ^e nicht bestimmbar ^e	
Interaktionstest	nicht bestimmbar ^c	
<p>a: HRs und <i>p</i>-Werte gelten für Fingolimod vs. Placebo oder IFN-β-1a i. m.; die <i>p</i>-Werte für die HRs wurden mittels des proportionalen Risikomodells nach Cox berechnet, adjustiert nach Behandlung.</p> <p>b: Der Interaktionstest (Interaktion Behandlungsgruppe/Geschlecht und Behandlungsgruppe/EDSS) wurde mithilfe des proportionalen Risikomodells nach Cox berechnet, nach Behandlungsgruppe, Geschlecht und EDSS adjustiert.</p> <p>c: Wegen kleiner Patientenzahlen in den Subpopulationen konnten keine komplexen Interaktionsmodelle durchgeführt werden.</p> <p>d: Wegen kleiner Patientenzahlen in den Subpopulationen konnten keine Schätzer berechnet werden.</p> <p>e: Kein Fall von Behinderungsprogression.</p>		

Tabelle 4-45: Ergebnisse für „Mittlere Änderung des EDSS bzw. MSFC“ aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS	Mittlere Änderung EDSS	Mittlere Änderung MSFC
Subgruppenanalyse Geschlecht^a		
Interaktionstest ITT	p=0,062	p=0,025
Interaktionstest IFN Non-Responder Def I	p=0,349	p=0,258
Interaktionstest IFN Non-Responder Def II	p=0,488	p=0,039
Interaktionstest Pat. mit schnell fortschr. schwerer RRMS	p<0,001	p=0,818
Subgruppenanalyse Krankheitsschwere^a		
Interaktionstest ITT	p=0,262	p=0,111
Interaktionstest IFN Non-Responder Def I	p=0,645	p=0,003
Interaktionstest IFN Non-Responder Def II	p=0,924	p=0,014
Interaktionstest Pat. mit schnell fortschr. schwerer RRMS	p=0,942	p=0,182
TRANSFORMS	Mittlere Änderung EDSS	Mittlere Änderung MSFC
Subgruppenanalyse Geschlecht^a		
Interaktionstest ITT	p=0,634	p=0,591
Interaktionstest IFN Non-Responder Def I	p=0,679	p=0,502
Interaktionstest IFN Non-Responder Def II	p=0,668	p=0,296
Interaktionstest Pat. mit schnell fortschr. schwerer RRMS	p=0,342	p=0,451
Subgruppenanalyse Krankheitsschwere^a		
Interaktionstest ITT	p=0,555	p=0,982
Interaktionstest IFN Non-Responder Def I	p=0,884	p=0,614
Interaktionstest IFN Non-Responder Def II	p=0,871	p=0,705
Interaktionstest Pat. mit schnell fortschr. schwerer RRMS	p=0,547	p=0,589
a: p-Werte für den Interaktionstest (Interaktion Behandlungsgruppe/Geschlecht und Behandlungsgruppe/EDSS) wurden mithilfe des Rang-ANCOVA Regressionsmodells berechnet, nach Behandlungsgruppe, Geschlecht und EDSS adjustiert.		

Die Interaktionstests und Subgruppenanalysen, die für beide Studien durchgeführt wurden, geben keinen konsistenten Hinweis auf einen Einfluss des Geschlechts oder der Krankheitsschwere (definiert über den Baseline EDSS) auf die Wirksamkeit von Fingolimod hinsichtlich der Reduktion des Anteils behinderungsprogressionsfreier Patienten.

Die Interaktionstests hinsichtlich der mittleren Änderung des EDSS beziehungsweise des MSFC zeigten ebenfalls keinen Einfluss von Geschlecht oder Krankheitsschwere. Die signifikanten p-Werte sind aufgrund der jeweiligen Gruppengrößen nicht interpretierbar. Die Tabellen mit den dazugehörigen Subgruppenauswertungen sind in Anhang 4-H zu finden (HTA-Tabellen 11.1.1 bis 12.4.8). Aus Gründen der Übersichtlichkeit wurden diese nicht im Hauptteil dargestellt.

Bei den beiden vorliegenden Phase III-Studien handelt es sich um verschiedene Studientypen (aktiv- beziehungsweise Placebo-kontrolliert), daher erscheint eine Zusammenfassung mittels Meta-Analyse wenig sinnvoll.

Endpunkt „gesundheitsbezogene Lebensqualität“

Tabelle 4-46: Operationalisierung von „gesundheitsbezogene Lebensqualität“

Studie	Operationalisierung
FREEDOMS	Mittlere Änderung des EQ-5D ^a zu Monat 24 gegenüber Baseline
TRANSFORMS	Mittlere Änderung des EQ-5D ^a zu Monat 12 gegenüber Baseline

a: Der Fragebogen EQ-5D ist ein Messinstrument für die gesundheitsbezogene Lebensqualität. Der EQ-5D ist zweigeteilt. Im ersten Teil wird die Lebensqualität über 5 Dimensionen mit jeweils 3 Antwortmöglichkeiten operationalisiert. Die daraus ermittelten Gesundheitszustände werden in einen Lebensqualitätsindex (utility score) zwischen 0 (Tod) und 1 (bestmöglicher Gesundheitszustand) transformiert. Im zweiten Teil des EQ-5D bewerten die Patienten ihren Gesundheitszustand auf einer visuellen Analogskala (VAS).

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-47: Bewertung des Verzerrungspotenzials für „gesundheitsbezogene Lebensqualität“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
EQ-5D						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial für den Endpunkt „gesundheitsbezogene Lebensqualität“ (EQ-5D) für beide Studien als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in beiden Fällen als „niedrig“ eingestuft wurde, der

Endpunkterheber (Patient) verblindet war, das ITT-Prinzip adäquat umgesetzt wurde und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte gab.

Tabelle 4-48: Ergebnisse für „Mittlere Änderung des EQ-5D (Index und VAS)“ aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)	
FREEDOMS	Mittlere Änderung EQ-5D (Index)	Mittlere Änderung EQ-5D (VAS)
Subgruppenanalyse Geschlecht^a		
Interaktionstest ITT	p=0,217	p=0,627
Interaktionstest IFN Non-Responder Def I	p=0,771	p=0,955
Interaktionstest IFN Non-Responder Def II	p=0,829	p=0,952
Interaktionstest Pat. mit schnell fortschr. schwerer RRMS	p=0,004	p=0,398
Subgruppenanalyse Krankheitsschwere^a		
Interaktionstest ITT	p=0,416	p=0,136
Interaktionstest IFN Non-Responder Def I	p=0,976	p=0,721
Interaktionstest IFN Non-Responder Def II	p=0,724	p=0,611
Interaktionstest Pat. mit schnell fortschr. schwerer RRMS	p=0,052	p=0,227
TRANSFORMS		
	Mittlere Änderung EQ-5D (Index)	Mittlere Änderung EQ-5D (VAS)
Subgruppenanalyse Geschlecht^a		
Interaktionstest ITT	p=0,200	p=0,918
Interaktionstest IFN Non-Responder Def I	p=0,441	p=0,658
Interaktionstest IFN Non-Responder Def II	p=0,533	p=0,314
Interaktionstest Pat. mit schnell fortschr. schwerer RRMS	p=0,167	p=0,037
Subgruppenanalyse Krankheitsschwere^a		
Interaktionstest ITT	p=0,311	p=0,215
Interaktionstest IFN Non-Responder Def I	p=0,959	p=0,581
Interaktionstest IFN Non-Responder Def II	p=0,925	p=0,539
Interaktionstest Pat. mit schnell fortschr. schwerer RRMS	p=0,375	p=0,123
a: p-Werte für den Interaktionstest (Interaktion Behandlungsgruppe/Geschlecht und Behandlungsgruppe/EDSS) wurden mithilfe des Rang-ANCOVA Regressionsmodells berechnet, nach Behandlungsgruppe, Geschlecht und EDSS adjustiert.		

Die Interaktionstests hinsichtlich der mittleren Änderung des EQ-5D (Index und VAS) zeigten keinen Einfluss von Geschlecht oder Krankheitsschwere. Die signifikanten p-Werte sind aufgrund der jeweiligen Gruppengrößen nicht interpretierbar. Die Tabellen mit den dazugehörigen Subgruppenauswertungen sind in Anhang 4-H zu finden (HTA-Tabellen

13.1.1 bis 14.4.8). Aus Gründen der Übersichtlichkeit wurden diese nicht im Hauptteil dargestellt.

Bei den beiden vorliegenden Phase III-Studien handelt es sich um verschiedene Studientypen (aktiv- beziehungsweise Placebo-kontrolliert), daher erscheint eine Zusammenfassung mittels Meta-Analyse wenig sinnvoll.

4.3.1.3.3 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien

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

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

Fingolimod ist die erste und bislang einzige Therapieoption zur Behandlung der RRMS, die mit Klasse-1-Evidenz nicht nur einen signifikanten medizinischen Nutzen gegenüber Placebo, sondern darüber hinaus auch einen signifikanten medizinischen Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie IFN- β -1a i. m. unter Beweis gestellt hat. Dieses Ergebnis der Head-to-Head-Studie TRANSFORMS verdeutlicht zum einen den Zusatznutzen von Fingolimod gegenüber einem Vertreter der Substanzklasse IFN- β hinsichtlich verschiedener wesentlicher klinischer Endpunkte (siehe nachstehend). Zum anderen markiert die Entwicklung von Fingolimod eine neue Qualität in der zielgerichteten Behandlung der RRMS. IFN- β wie auch Glatirameracetat stellen in der Historie der RRMS und ihrer krankheitsmodifizierenden Beeinflussbarkeit sicherlich entscheidende Meilensteine dar. Bei aller hinreichend gut dokumentierten Wirksamkeit ist diesen beiden Therapieansätzen jedoch gemeinsam, dass ihre Wirkprinzipien im Grunde genommen bis heute nicht verstanden sind. Bei Fingolimod hingegen handelt es sich erstmals um eine sogenannte zielgerichtete Therapie mit klar definiertem Wirkansatz (ausführlich beschrieben in Modul 2), deren immunologische Effekte nach Absetzen vollständig reversibel und deshalb sehr gut steuerbar sind.

Das Studienprogramm zur Bewertung des medizinischen Nutzens und Zusatznutzens sowie der Sicherheit von Fingolimod bei RRMS ist außerordentlich umfangreich und stellt die bislang größte Datensammlung auf dem Gebiet der RRMS für ein Einzelpräparat dar. Es liefert Daten von mehr als 4.000 Patienten und umfasst eine Wirkstoff-Exposition von mehr als 10.000 Patientenjahren. Damit gibt das Studienprogramm nicht nur detaillierten Aufschluss über den medizinischen Nutzen und Zusatznutzen des Therapieansatzes, sondern erlaubt auch eine fundierte Einschätzung der Sicherheit in der Anwendung des Präparates.

Bewertung des medizinischen Nutzens:

Als Ergebnis des Placebovergleichs lässt sich ein medizinischer Nutzen von Fingolimod feststellen. Der medizinische Nutzen von Fingolimod gegenüber Placebo zur Behandlung der RRMS beziehungsweise im Speziellen zur Behandlung hochaktiver Formen der MS wird im Folgenden zusammengefasst (FREEDOMS-Studie, berücksichtigt wird nur die zugelassene Dosierung Fingolimod 0,5 mg 1 x täglich; IFN Non-Responder Def I und Def II umfassen jeweils die Anwendungsgebiete a1 und a2, Patienten mit schnell fortschreitender schwerer RRMS entsprechen Anwendungsgebiet b). Generell müssen die Auswertungen in den Subpopulationen aufgrund der reduzierten Fallzahlen mit Vorsicht interpretiert werden.

Schubbezogene Parameter

Fingolimod

- reduziert die jährliche Schubrate um 54 % im Vergleich zu Placebo in der Gesamtpopulation (ARR ratio (95 %-KI): 0,462 (0,373; 0,571); $p < 0,001$), um 62 % bzw. 71 % bei den IFN-Non-Respondern (Def I und Def II) (0,379 (0,212; 0,677); $p = 0,001$ und 0,287 (0,158; 0,524); $p < 0,001$) und um 67 % bei Patienten mit schnell fortschreitender schwerer RRMS (0,329 (0,175; 0,620); $p < 0,001$).
- reduziert das Schubrisiko um 52 % im Vergleich zu Placebo in der Gesamtpopulation (HR (95 %-KI): 0,48 (0,39; 0,61); $p < 0,001$), um 60 % bzw. 71 % bei den IFN Non-Respondern (Def I und Def II) (0,40 (0,22; 0,73); $p = 0,003$ und 0,29 (0,15; 0,53); $p < 0,001$) und um 68 % bei Patienten mit schnell fortschreitender schwerer RRMS (0,32 (0,16; 0,61); $p < 0,001$).
- reduziert die Anzahl der schweren Schübe im Vergleich zu Placebo in der Gesamtpopulation: Anzahl (%) der Schübe - schwere Schübe 10 (5,8 %) vs. 35 (9,7 %); mittelschwere Schübe 102 (59,3 %) vs. 205 (57,1 %); milde Schübe 60 (34,9 %) vs. 119 (33,1 %); $p = 0,322$ (overall).

die Anzahl der schweren Schübe im Vergleich zu Placebo bei den IFN Non-Respondern (Def I und Def II): Anzahl (%) der Schübe - schwere Schübe 1 (4,0 %) vs. 7 (13,0 %) und 1 (4,8 %) vs. 7 (12,5 %); mittelschwere Schübe 14 (56,0 %) vs. 25 (46,3 %) und 11 (52,4 %) vs. 29 (51,8 %); milde Schübe 10 (40,0 %) vs. 22 (40,7 %) und 9 (42,9 %) vs. 20 (35,7 %); $p = 0,538$ (overall) und $p = 0,660$ (overall)

die Anzahl der schweren Schübe im Vergleich zu Placebo bei Patienten mit schnell fortschreitender schwerer RRMS: Anzahl (%) der Schübe - schwere Schübe 1 (4,5 %) vs. 4 (8,7 %); mittelschwere Schübe 17 (77,3 %) vs. 32 (69,6 %); milde Schübe 4 (18,2 %) vs. 10 (21,7 %); $p = 0,833$ (overall)

- reduziert die Anzahl der krankenhauspflchtigen Schübe im Vergleich zu Placebo: Anzahl (%) der Schübe in der Gesamtpopulation: 63 (36,6 %) vs. 146 (40,7 %); $p = 0,394$; bei den IFN Non-Respondern (Def I und Def II): 8 (32,0 %) vs. 17 (31,5 %) und 8 (38,1 %) vs. 17

(30,4 %); $p=1,000$ und $0,589$; bei Patienten mit schnell fortschreitender schwerer RRMS: 6 (27,3 %) vs. 19 (41,3 %); $p=0,296$.

- reduziert die Anzahl der steroidpflichtigen Schübe im Vergleich zu Placebo: Anzahl (%) der Schübe in der Gesamtpopulation: 140 (81,4 %) vs. 303 (84,4 %); $p=0,385$; bei den IFN Non-Respondern (Def I und Def II): 16 (64,0 %) vs. 45 (83,3 %) und 14 (66,7 %) vs. 49 (87,5 %); $p=0,083$ und $0,048$; bei Patienten mit schnell fortschreitender schwerer RRMS: 21 (95,5 %) vs. 37 (80,4 %); $p=0,149$.

Behinderungsbezogene Parameter

- Fingolimod senkt das Risiko der Behinderungsprogression im Vergleich zu Placebo in der Gesamtpopulation signifikant (HR (95 %-KI): 0,70 (0,52; 0,96); $p=0,024$); und tendenziell bei den IFN-Non-Respondern (Def I und Def II) (0,68 (0,29; 1,62), $p=0,386$ und 0,64 (0,27; 1,51); $p=0,306$) und bei Patienten mit schnell fortschreitender schwerer RRMS (0,73 (0,25; 2,07); $p=0,548$).
- Unter Fingolimod verschlechtern sich die Behinderungsparameter nicht beziehungsweise verbessern sich leicht im Vergleich zu Placebo:

In der Gesamtpopulation: Mittlere (\pm SEM) Veränderung des EDSS $0,00 \pm 0,045$ vs. $0,13 \pm 0,051$; $p=0,004$ und des MSFC $0,03 \pm 0,021$ vs. $-0,06 \pm 0,032$; $p=0,009$.

IFN Non-Responder Def I: Mittlere (\pm SEM) Veränderung des EDSS $0,04 \pm 0,142$ vs. $-0,15 \pm 0,163$; $p=0,852$ und des MSFC $-0,01 \pm 0,055$ vs. $-0,24 \pm 0,094$; $p=0,144$.

IFN non-responder Def II: Mittlere (\pm SEM) Veränderung des EDSS $-0,03 \pm 0,139$ vs. $-0,07 \pm 0,164$; $p=0,736$ und des MSFC $-0,05 \pm 0,057$ vs. $-0,27 \pm 0,095$; $p=0,083$.

Patienten mit schnell fortschreitender schwerer RRMS: Mittlere (\pm SEM) Veränderung des EDSS $-0,04 \pm 0,148$ vs. $-0,04 \pm 0,211$; $p=0,818$ und des MSFC $0,07 \pm 0,083$ vs. $-0,13 \pm 0,134$; $p=0,199$.

Gesundheitsbezogene Lebensqualität

- Unter Fingolimod bleibt die gesundheitsbezogene Lebensqualität konstant beziehungsweise verbessert sich leicht im Vergleich zu Placebo:
- In der Gesamtpopulation: Mittlere (\pm SEM) Veränderung des EQ-5D Index $0,00 \pm 0,009$ vs. $0,01 \pm 0,011$; $p=0,419$ und der EQ-5D VAS $-0,23 \pm 0,817$ vs. $-1,04 \pm 0,943$; $p=0,834$

IFN Non-Responder Def I: Mittlere (\pm SEM) Veränderung des EQ-5D Index $0,00 \pm 0,032$ vs. $0,05 \pm 0,031$; $p=0,523$ und der EQ-5D VAS $1,02 \pm 2,465$ vs. $2,38 \pm 3,004$; $p=0,907$

IFN Non-Responder Def II: Mittlere (\pm SEM) Veränderung des EQ-5D Index $-0,01 \pm 0,031$ vs. $0,00 \pm 0,020$; $p=0,896$ und der EQ-5D VAS $1,31 \pm 2,786$ vs. $1,00 \pm 3,254$; $p=0,730$

Patienten mit schnell fortschreitender schwerer RRMS: Mittlere (\pm SEM) Veränderung des EQ-5D Index $0,03 \pm 0,026$ vs. $-0,07 \pm 0,041$; $p=0,037$ und der EQ-5D VAS $2,08 \pm 2,473$ vs. $-2,96 \pm 4,547$; $p=0,251$

Sicherheit

- Die Gesamtinzidenz von UE (RR 1,02; 95 %-KI (0,98; 1,06)), SUE (RR 0,78; 95 %-KI (0,53; 1,14)) Behandlungsabbrüchen aufgrund von UE (RR 1,02; 95 %-KI (0,63; 1,63)) ist für Fingolimod und Placebo vergleichbar.
- Ein erhöhtes Risiko gegenüber Placebo lässt sich für Lymphopenien (RR 5,90; 95 %-KI (1,33; 26,21)), Leukopenien (RR 9,84, 95 %-KI (1,26; 76,49)) und Leberwerterhöhungen (RR 2,73, 95 %-KI (1,80; 4,15)) ableiten. Die berichteten Ereignisse waren nach Absetzen der Medikation reversibel.

Bewertung des medizinischen Zusatznutzens:

Als Ergebnis des direkten Wirksamkeitsvergleichs lässt sich ein medizinischer Zusatznutzen von Fingolimod gegenüber der zweckmäßigen Vergleichstherapie IFN- β -1a i. m. feststellen. Dieser wird im Folgenden zusammengefasst (TRANSFORMS-Studie, berücksichtigt wird nur die zugelassene Dosierung Fingolimod 0,5 mg 1 x täglich; IFN Non-Responder Def I und Def II umfassen jeweils die Anwendungsgebiete a1 und a2, Patienten mit schnell fortschreitender schwerer RRMS entsprechen Anwendungsgebiet b). Generell müssen die Auswertungen in den Subpopulationen aufgrund der reduzierten Fallzahlen mit Vorsicht interpretiert werden.

Schubbezogene Parameter

- Fingolimod reduziert die jährliche Schubrate um 52 % im Vergleich zu IFN- β -1a i. m. in der Gesamtpopulation (ARR ratio (95 %-KI) 0,484 (0,367; 0,639); $p<0,001$), um 61 % bei den IFN Non-Respondern (Def I und Def II) (0,391 (0,246; 0,623); $p<0,001$ und 0,389 (0,244; 0,619); $p<0,001$) und um 25 % bei Patienten mit schnell fortschreitender MS (0,746 (0,238; 2,333); $p=0,614$).

- Fingolimod reduziert das Schubrisiko um 48 % im Vergleich zu IFN- β -1a i. m. in der Gesamtpopulation (HR (95 %-KI): 0,52 (0,39; 0,69), $p < 0,001$), um 62 % bzw 58 % bei den IFN-Non-Respondern (Def I und II) (0,38 (0,24; 0,61); $p < 0,001$ und 0,42 (0,26; 0,67), $p < 0,001$) und um 24 % bei Patienten mit schnell fortschreitender schwerer RRMS (0,76 (0,24; 2,39); $p = 0,637$).
- Fingolimod reduziert die Anzahl der schweren Schübe im Vergleich zu IFN- β -1a i. m. in der Gesamtpopulation: Anzahl (%) der Schübe - schwere Schübe 10 (11,2 %) vs. 30 (16,8 %); mittelschwere Schübe 48 (53,9 %) vs. 105 (58,7 %); milde Schübe 31 (34,8 %) vs. 44 (24,6 %); $p = 0,166$ (overall).

die Anzahl der schweren Schübe im Vergleich zu IFN- β -1a i. m. bei den IFN Non-Respondern (Def I und II): Anzahl (%) der Schübe - schwere Schübe 6 (18,8 %) vs. 17 (23,6 %) und 5 (15,2 %) vs. 17 (25,4 %); mittelschwere Schübe 20 (62,5 %) vs. 37 (51,4 %) und 20 (60,6 %) vs. 35 (52,2 %); milde Schübe 6 (18,8 %) vs. 18 (25,0 %) und 8 (24,2 %) vs 15 (22,4 %); $p = 0,656$ (overall, Def I) und $p = 0,552$ (overall, Def II)

die Anzahl der schweren Schübe im Vergleich zu IFN- β -1a i. m. bei Patienten mit schnell fortschreitender schwerer RRMS: Anzahl (%) der Schübe - schwere Schübe 0 (0,0 %) vs. 1 (11,1 %); mittelschwere Schübe 4 (66,7 %) vs. 5 (55,6 %); milde Schübe 2 (33,3 %) vs. 3 (33,3 %); $p = 1,000$ (overall)

- Fingolimod reduziert die Anzahl der krankenhauspflchtigen Schübe im Vergleich zu IFN- β -1a i. m: Anzahl (%) der Schübe in der Gesamtpopulation: 11 (12,4 %) vs. 36 (20,1 %); $p = 0,128$; bei den IFN Non-Respondern (Def I und II): 4 (12,5 %) vs. 14 (19,4 %) und 4 (12,1 %) vs. 11 (16,4 %); $p = 0,575$ und $p = 0,768$; bei Patienten mit schnell fortschreitender schwerer RRMS: 0 (0,0 %) vs. 1 (11,1 %); $p = 1,000$.
- Fingolimod reduziert die Anzahl der steroidpflichtigen Schübe im Vergleich zu IFN- β -1a i. m: Anzahl (%) der Schübe in der Gesamtpopulation: 65 (73,0 %) vs. 148 (82,7 %); $p = 0,078$; bei den IFN Non-Respondern (Def I und II): 26 (81,3 %) vs. 63 (87,5 %) und 24 (72,7 %) vs. 57 (85,1 %); $p = 0,546$ und $p = 0,177$.; bei Patienten mit schnell fortschreitender schwerer RRMS: 3 (50,0 %) vs. 6 (66,7 %); $p = 0,622$.

Behinderungsbezogene Parameter

- Fingolimod senkt das Risiko der Behinderungsprogression im Vergleich zu IFN- β -1a i. m in der Gesamtpopulation tendenziell (HR (95 %-KI): 0,71 (0,42; 1,21), $p = 0,209$); und tendenziell bei den IFN-Non-Respondern (Def I und II) (0,60 (0,26; 1,41); $p = 0,244$ und 0,58 (0,23; 1,44); $p = 0,242$). Bei Patienten mit schnell fortschreitender schwerer RRMS konnte keine Hazard Ratio ermittelt werden (kein bestätigter Fall von Behinderungsprogression für die Fingolimod-Gruppe).

- Unter Fingolimod verschlechtern sich die Behinderungsparameter nicht beziehungsweise verbessern sich leicht im Vergleich zu IFN- β -1a i. m.:

In der Gesamtpopulation: Mittlere (\pm SEM) Veränderung des EDSS $-0,08 \pm 0,040$ vs. $0,01 \pm 0,092$; $p=0,059$ und des MSFC $0,04 \pm 0,021$ vs. $-0,03 \pm 0,025$; $p=0,025$.

IFN Non-Responder Def I: Mittlere (\pm SEM) Veränderung des EDSS $-0,09 \pm 0,063$ vs. $-0,04 \pm 0,073$; $p=0,116$ und des MSFC $0,02 \pm 0,028$ vs. $-0,13 \pm 0,057$; $p=0,006$.

IFN non-responder Def II: Mittlere (\pm SEM) Veränderung des EDSS $-0,11 \pm 0,060$ vs. $0,04 \pm 0,079$; $p=0,074$ und des MSFC $-0,01 \pm 0,030$ vs. $-0,09 \pm 0,041$; $p=0,010$.

Patienten mit schnell fortschreitender schwerer RRMS: Mittlere (\pm SEM) Veränderung des EDSS $-0,06 \pm 0,063$ vs. $0,17 \pm 0,139$; $p=0,098$ und des MSFC $0,02 \pm 0,033$ vs. $0,03 \pm 0,061$; $p=0,856$.

Gesundheitsbezogene Lebensqualität

- Unter Fingolimod bleibt die gesundheitsbezogene Lebensqualität konstant beziehungsweise verbessert sich leicht im Vergleich zu IFN- β -1a i. m.:
- In der Gesamtpopulation: Mittlere (\pm SEM) Veränderung des EQ-5D Index $0,01 \pm 0,009$ vs. $-0,00 \pm 0,010$; $p=0,100$ und der EQ-5D VAS $1,89 \pm 0,764$ vs. $0,63 \pm 0,774$; $p=0,257$

IFN Non-Responder Def I: Mittlere (\pm SEM) Veränderung des EQ-5D Index $0,01 \pm 0,015$ vs. $0,03 \pm 0,015$; $p=0,844$ und der EQ-5D VAS $3,22 \pm 1,160$ vs. $-0,92 \pm 1,344$; $p=0,042$

IFN Non-Responder Def II: Mittlere (\pm SEM) Veränderung des EQ-5D Index $-0,01 \pm 0,014$ vs. $0,02 \pm 0,017$; $p=0,624$ und der EQ-5D VAS $2,95 \pm 1,091$ vs. $-1,40 \pm 1,441$; $p=0,029$

Patienten mit schnell fortschreitender schwerer RRMS: Mittlere (\pm SEM) Veränderung des EQ-5D Index $0,04 \pm 0,023$ vs. $-0,02 \pm 0,029$; $p=0,101$ und der EQ-5D VAS $2,78 \pm 4,770$ vs. $-0,00 \pm 3,006$; $p=0,699$

Sicherheit

- Die Gesamtinzidenz von UE (RR 0,94; 95 %-KI (0,89; 0,98)), SUE (RR 1,26; 95 %-KI (0,75; 2,11)) Behandlungsabbrüchen aufgrund von UE (RR 1,51; 95 %-KI (0,81; 2,80)) ist für Fingolimod und IFN- β -1a i. m. vergleichbar.
- Ein erhöhtes Risiko gegenüber IFN- β -1a i. m. lässt sich für Leberwerterhöhungen (RR 3,60; 95 %-KI (2,14; 6,07)) ableiten. Die berichteten Ereignisse waren nach Absetzen der

Medikation reversibel. Grippeartige Symptome kommen unter Fingolimod im Gegensatz zu IFN- β -1a i. m. so gut wie nicht vor (RR 0,09; 95 %-KI (0,06; 0,16)).

Fazit zu schubbezogenen Parametern

Die Senkung der jährlichen Schubraten durch Fingolimod vs. IFN- β -1a i. m. fällt deutlicher aus als die früher dokumentierten Schubratensenkungen unter immunmodulierender Basistherapie vs. Placebo, die kaum über 30 % hinausgingen (39-41).

Bei der Vergleichssubstanz für Fingolimod in TRANSFORMS handelt es sich um die zweckmäßige Vergleichstherapie IFN- β -1a i. m. in der Behandlung der RRMS. Damit erbringt Fingolimod einen Zusatznutzen gegenüber den herkömmlichen Therapieoptionen, die die bislang erreichten Effekte hinsichtlich der klinisch sichtbaren Anzeichen für ein Fortbestehen der Krankheitsaktivität deutlich übertrifft. Dies wird besonders deutlich in dem klar gesteigerten Anteil von Patienten, die frei von Schubaktivität sind. Die Verringerung der Schubaktivität ist insofern substanziell, als sie sich ausschließlich auf Schübe mittleren bis hohen Schweregrades bezieht.

Der medizinische Nutzen und Zusatznutzen von Fingolimod gegenüber der zweckmäßigen Vergleichstherapie IFN- β 1a i. m. erweist sich als nachhaltig und über einen Zeitraum von zwei Jahren gesichert.

Fazit zu behinderungsbezogenen Parametern

Der Anteil an Patienten ohne Behinderungsprogression war unter Fingolimod gegenüber den jeweiligen Vergleichsbehandlungen erhöht. Fingolimod zeigt damit einen medizinischen Nutzen und Zusatznutzen hinsichtlich der Behinderungsprogression. Die mittleren EDSS-Werte und MSFC z-Punktwerte blieben unter Fingolimod-Behandlung stabil oder verbesserten sich leicht, während sie sich in den Gruppen mit Placebo und IFN- β -1a i. m. jedoch leicht verschlechterten. Sowohl in TRANSFORMS als auch in FREEDOMS sprachen die EDSS-Werte und MSFC z-Punktwerte für Fingolimod 0,5 mg im Vergleich zu den jeweiligen Vergleichsbehandlungen.

Fazit zur gesundheitsbezogenen Lebensqualität

Tendenziell zeigt sich Fingolimod auch hinsichtlich der gesundheitsbezogenen Lebensqualität überlegen, die Lebensqualität konnte im Gegensatz zu Placebo oder IFN- β -1a i. m. konstant gehalten werden bzw. tendenziell gesteigert werden. Diese Ergebnisse sind aber aufgrund der geringen Spezifität des EQ-5D (indikationsübergreifendes Instrument) eher von untergeordneter Aussagekraft. Zudem konnte keine Aussage getroffen werden, inwiefern eine

orale Therapie im Gegensatz zu einer Injektionstherapie hinsichtlich der Lebensqualität von Vorteil ist. Das Double-Dummy Design der TRANSFORMS-Studie lässt das nicht zu, da jeder Patient sowohl eine Kapsel als auch eine Injektion erhalten hat.

Fazit zum Sicherheitsprofil von Fingolimod

Fingolimod wird im Allgemeinen von Patienten mit schubförmig verlaufender MS gut vertragen. Die Inzidenz unerwünschter Ereignisse in den Studien war zwischen den Behandlungsgruppen vergleichbar und die meisten Ereignisse waren von leichter bis mäßiger Intensität. Die Inzidenz schwerwiegender unerwünschter Ereignisse war unter Fingolimod-Behandlung niedrig. Die Gesamtinzidenz von Infektionen war in den Behandlungsgruppen ähnlich und die meisten Infektionen waren von leichter bis mäßiger Intensität. Wirkungen von Fingolimod auf das Herz-Kreislauf-System sind vorübergehender Natur und klinisch gutartig. Unter Fingolimod-Behandlung kann ein erhöhtes Risiko für Makulaödeme beobachtet werden, welches nach Absetzen der Behandlung reversibel ist. Klinische Studien liefern keine Evidenz für einen Anstieg maligner Erkrankungen der Haut, Brust oder anderer Organe unter Behandlung mit Fingolimod.

Die TRANSFORMS-Daten erlauben eine vergleichende Bewertung der Verträglichkeit von Fingolimod 0,5 mg und IFN- β -1a i. m. Grippeähnliche Beschwerden traten mit einer Häufigkeit von 36,9 % unter IFN- β -1a i. m. auf. Unter Fingolimod kam es häufiger zu einem Anstieg von Leberenzymwerten.

Ein erhöhtes Risiko opportunistischer Infektionen unter der Therapie mit Fingolimod ist nach aktueller Datenlage nicht gegeben. In den Fingolimod-Zulassungsstudien (FREEDOMS, TRANSFORMS) wurden insgesamt 1.703 Patienten über einen Zeitraum von 12 beziehungsweise 24 Monaten behandelt (in der jetzt zugelassenen Dosierung von 0,5 mg, aber auch in der höheren Dosierung 1,25 mg). Im MS-Programm wurden Patienten mit insgesamt mehr als 10.000 Patientenjahren mit Fingolimod behandelt. In diesen Studien trat kein einziger Fall von progressiver multifokaler Leukenzephalopathie (PML) auf.

Diese Daten deuten darauf hin, dass das Risiko, eine PML unter Fingolimod zu entwickeln, nicht höher ist als im unbehandelten Teil der Bevölkerung.

Auswirkungen des Nebenwirkungsprofils auf die Therapieadhärenz

Mangelnde Adhärenz beziehungsweise Compliance ist ein großes Problem in der derzeitigen MS-Basistherapie. Die eingeschränkte Adhärenz hat sowohl direkte Konsequenzen für das Wohlergehen des einzelnen Patienten als auch für das Gesundheitssystem. Die in der Literatur am häufigsten angeführten Gründe für Therapieabbrüche sind dabei Nebenwirkungen und ein subjektiv wahrgenommener Mangel an Wirksamkeit. Es ist anzumerken, dass die Adhärenzraten in prospektiven, groß angelegten Studien häufig deutlich über den Adhärenzraten in der klinischen Praxis liegen (42;43).

Die Therapie mit derzeitig zugelassenen Basistherapeutika ist zwar sicher, die typischen systemischen (v. a. grippeähnliche Symptome) applikationsbedingten lokalen Nebenwirkungen (v. a. Reaktionen an der Einstichstelle) der zweckmäßigen Vergleichstherapie beeinflussen allerdings die Therapieadhärenz und somit die klinische Wirksamkeit (44). Unter Fingolimod bleiben diese generell aus, die Patienten erfahren somit deutlich weniger Einschränkungen im privaten wie im beruflichen Leben. Fingolimod bietet hier klare Vorteile gegenüber einer Therapie mit IFN- β -1a i. m. und erfüllt wesentliche Voraussetzungen für eine optimale Therapieadhärenz.

4.3.2 Weitere Unterlagen

4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

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

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu Studien für indirekte Vergleiche. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung. Benennen Sie

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

4.3.2.1.1.1 Studien des pharmazeutischen Unternehmers

Benennen Sie in der nachfolgenden Tabelle alle Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war.

Von den randomisierten kontrollierten Studien, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden, sollen die Studien, deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle randomisierten kontrollierten Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.

Benennen Sie nur randomisierte, kontrollierte Studien, 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“, Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer und Therapierme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-49: Liste der Studien des pharmazeutischen Unternehmers – Studien für indirekte Vergleiche

Studienbezeichnung ^a	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
FTY720D1201	ja (für Japan)	abgeschlossen	6 Monate	Fingolimod 0,5 mg Fingolimod 1,25 mg Placebo
FTY720D2201	nein	abgeschlossen	6 Monate	Fingolimod 1,25 mg Fingolimod 5,0 mg Placebo
FTY720D2301 FREEDOMS	ja	abgeschlossen	24 Monate	Fingolimod 0,5 mg Fingolimod 1,25 mg Placebo
FTY720D2302 TRANSFORMS	ja	abgeschlossen	12 Monate	Fingolimod 0,5 mg Fingolimod 1,25 mg IFN-β-1a i. m.
FTY720D2306 INFORMS	nein	laufend	36+ Monate	Fingolimod 1,25 mg ^b Placebo
FTY720D2309 FREEDOMS II	nein	laufend	24 Monate	Fingolimod 0,5 mg Fingolimod 1,25 mg ^b Placebo
FTY720D2320	nein	abgeschlossen ^c	3 Monate	Fingolimod 0,5 mg Placebo
FTY720DUS01	nein	laufend	6 Monate	Fingolimod 0,5 mg „Standard-of-Care“ ^c -DMT
CFTY720DIT01	nein	laufend	18 Monate	Fingolimod 0,5 mg IFN-β-1b s. c.
CFTY720DIT02	nein	laufend	6 Monate	Fingolimod 0,5 mg IFN-β-1a, IFN-β-1b oder Glatirameracetat
<p>a: Nicht gelistet sind Phase I-Studien</p> <p>b: Im November 2009 wurden alle Patienten auf Fingolimod 0,5 mg umgestellt</p> <p>c: noch keine Auswertung vorliegend</p> <p>Anmerkung: Eine Übersicht zu allen mit Fingolimod in der RRMS durchgeführten und laufenden Studien ist in Modul 5 hinterlegt (Modul5, Studienliste_Fingolimod.pdf).</p>				

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

Die Informationen in Tabelle 4-4 wurden zuletzt am 26. Juli 2011 aktualisiert, d. h. die Tabelle enthält den Studienstatus zu diesem Datum.

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

Tabelle 4-50: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – Studien für indirekte Vergleiche

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
FTY720D1201	Studiendauer zu kurz (6-Monats-Studie)
FTY720D2201	untersucht nicht die zugelassene Dosierung Studiendauer zu kurz (6-Monats-Studie)
FTY720D2306 INFORMS	laufende Studie, andere Indikation (PPMS)
FTY720D2309 FREEDOMS II	laufende Studie
FTY720D2320	noch keine Auswertung vorliegend
FTY720DUS01	laufende Studie
CFTY720DIT01	laufende Studie
CFTY720DIT02	laufende Studie

Stellen Sie die abgebrochenen und laufenden Studien, die nicht in die Bewertung eingehen, in Anhang 4-D und Anhang 4-E dar.

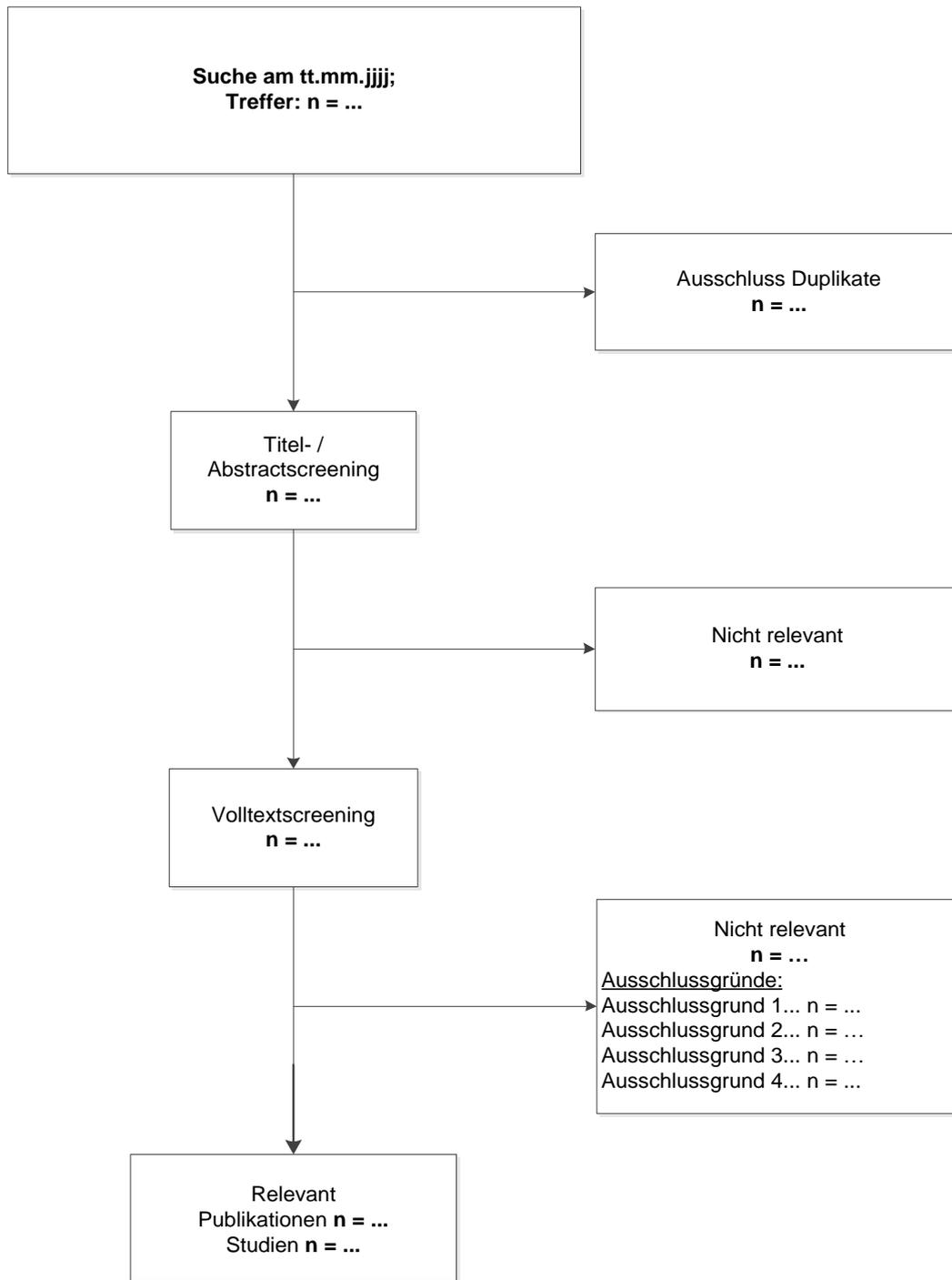
4.3.2.1.2 Studien aus der bibliografischen Literaturrecherche

Sofern eine bibliografische Literaturrecherche durchgeführt wurde, beschreiben Sie nachfolgend das Ergebnis dieser Recherche. 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 Studien unter Nennung des Ausschlussgrunds in Anhang 4-C.

Falls die Recherche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel gemeinsam mit der Recherche zu einer anderen Fragestellung (z. B. zu indirekten Vergleichen) durchgeführt wurde, unterteilen Sie die Angaben zu relevanten Treffern und Studien entsprechend diesen Fragestellungen.

[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 der Erstellung des Flussdiagramms an dem nachfolgenden Beispiel.



Es wurden drei separate Literaturrecherchen durchgeführt. Die Suche nach RCT mit dem zu bewertenden Arzneimittel Fingolimod wurden bereits in Abschnitt 4.3.1.1.2 beschrieben. Die Suchergebnisse gelten ebenso für den indirekten Vergleich.

Um Studien zu identifizieren, die für eine Netzwerkanalyse relevant sind, wurde nach RCT mit den verfügbaren Basistherapeutika gesucht, IFN- β -Präparate und Glatirameracetat. Die Suchkriterien sind in Anhang 4-A aufgeführt,

Die Suchen wurden jeweils geteilt nach Placebo- und aktiv-kontrollierten Studien. Die Suche nach aktiv-kontrollierten IFN- β -Studien war dabei definiert als Suche nach Studien, die IFN- β -Präparate untereinander vergleichen (IFN- β -1a i.m. oder s.c. vs. IFN- β -1b s.c. und IFN- β -1a s.c. vs. IFN- β -1a i.m.). Studien von IFN- β gegen Glatirameracetat wurden über die Suche nach aktiv-kontrollierten (ohne Einschränkung) Glatirameracetat-Studien identifiziert.

Die Selektionskriterien sind in Abschnitt 4.2.2 aufgeführt. Gemäß dieser Kriterien wurden keine Studien mit Glatirameracetat identifiziert, die die IFN Non-Responder definieren. Entweder war eine Vortherapie von vornherein ausgeschlossen, oder die Vortherapien konnten keine IFN beinhalten, weil die Studien vor Einführung der IFN-Präparate rekrutiert hatten. Ein indirekter Vergleich mit Bezug auf die für Fingolimod relevanten Populationen ist daher nicht durchführbar. Die nachfolgend dargestellte Netzwerkanalyse bezieht sich daher auf RRMS-Patienten allgemein und ist für die Fragestellung des Vergleichs von Fingolimod mit Glatirameracetat die best-verfügbare Evidenz.

Nachfolgend sind die Flussdiagramme für die Suche zu den IFN- β -Präparaten und Glatirameracetat dargestellt.

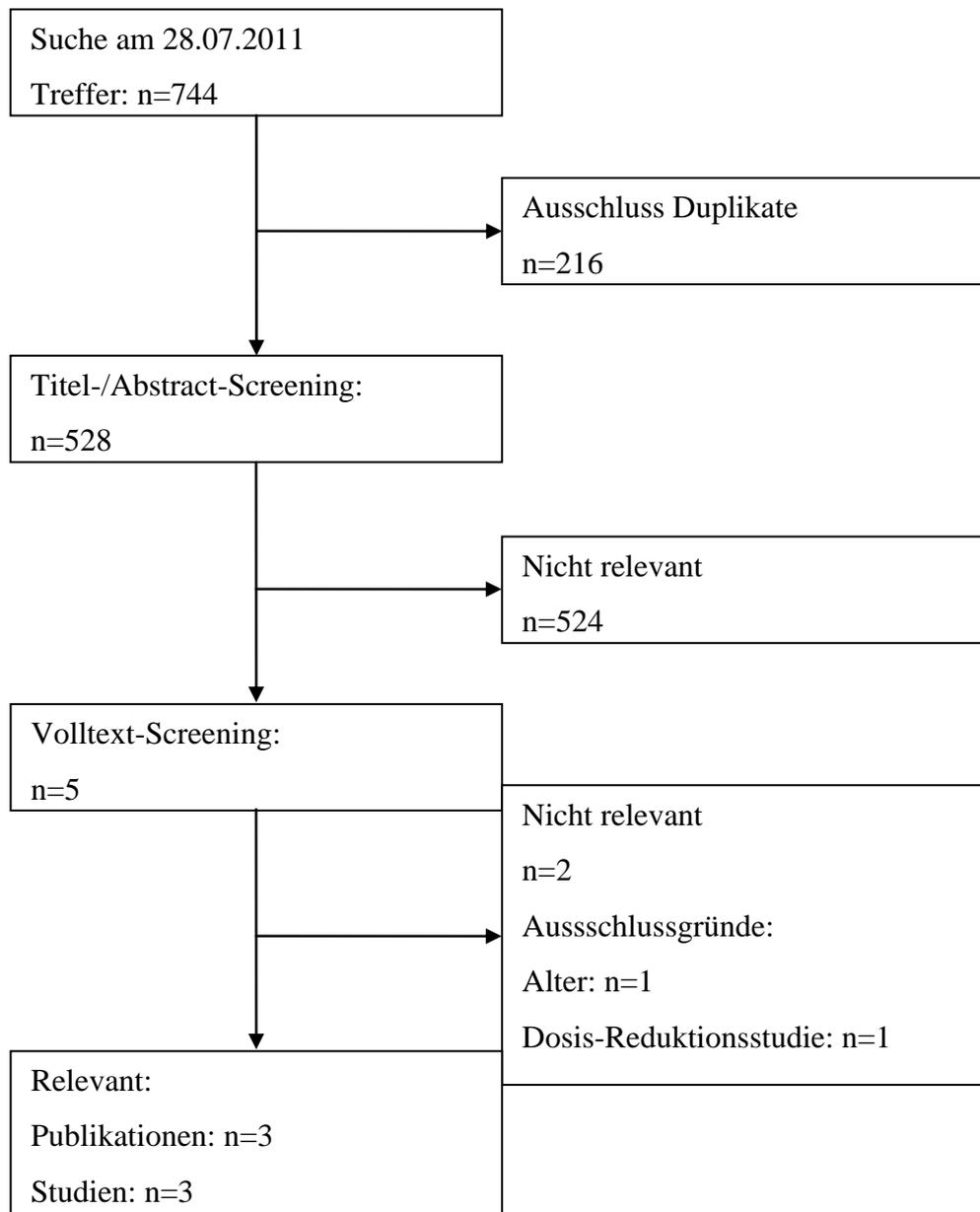
Abbildung 4-21: Flussdiagramm der bibliografischen Literaturrecherche zu IFN- β -Studien (aktiv-kontrolliert mit IFN- β -Präparaten; IFN- β -1a i. m./s. c. vs. IFN- β -1b s. c.)

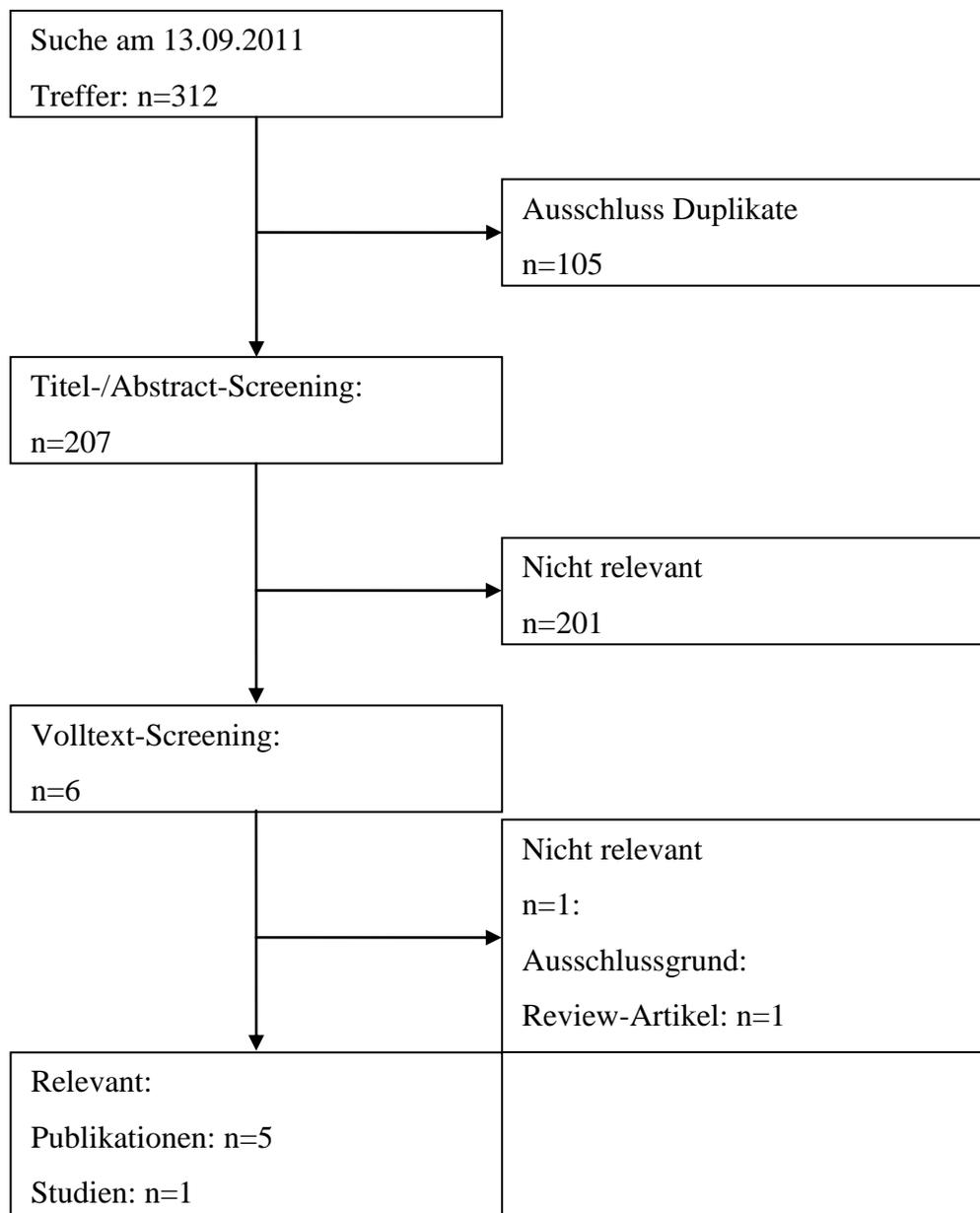
Abbildung 4-22: Flussdiagramm der bibliografischen Literaturrecherche zu IFN- β -Studien (aktiv-kontrolliert mit IFN- β -Präparaten; IFN- β -1a i. m vs. IFN- β -1a s. c.)

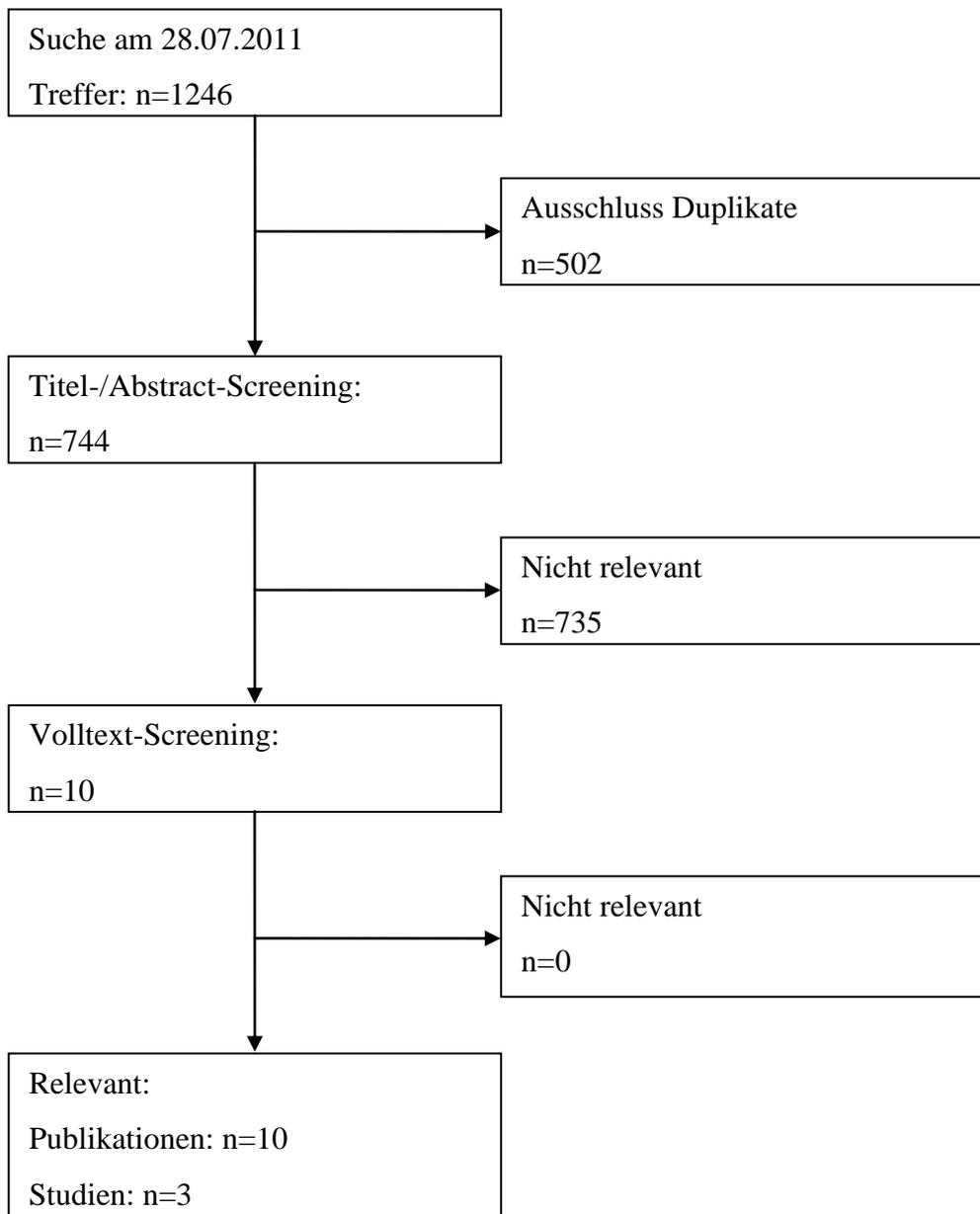
Abbildung 4-23: Flussdiagramm der bibliografischen Literaturrecherche zu IFN- β -Studien (Placebo-kontrolliert)

Abbildung 4-24: Flussdiagramm der bibliografischen Literaturrecherche zu Glatirameracetat-Studien (aktiv-kontrolliert)

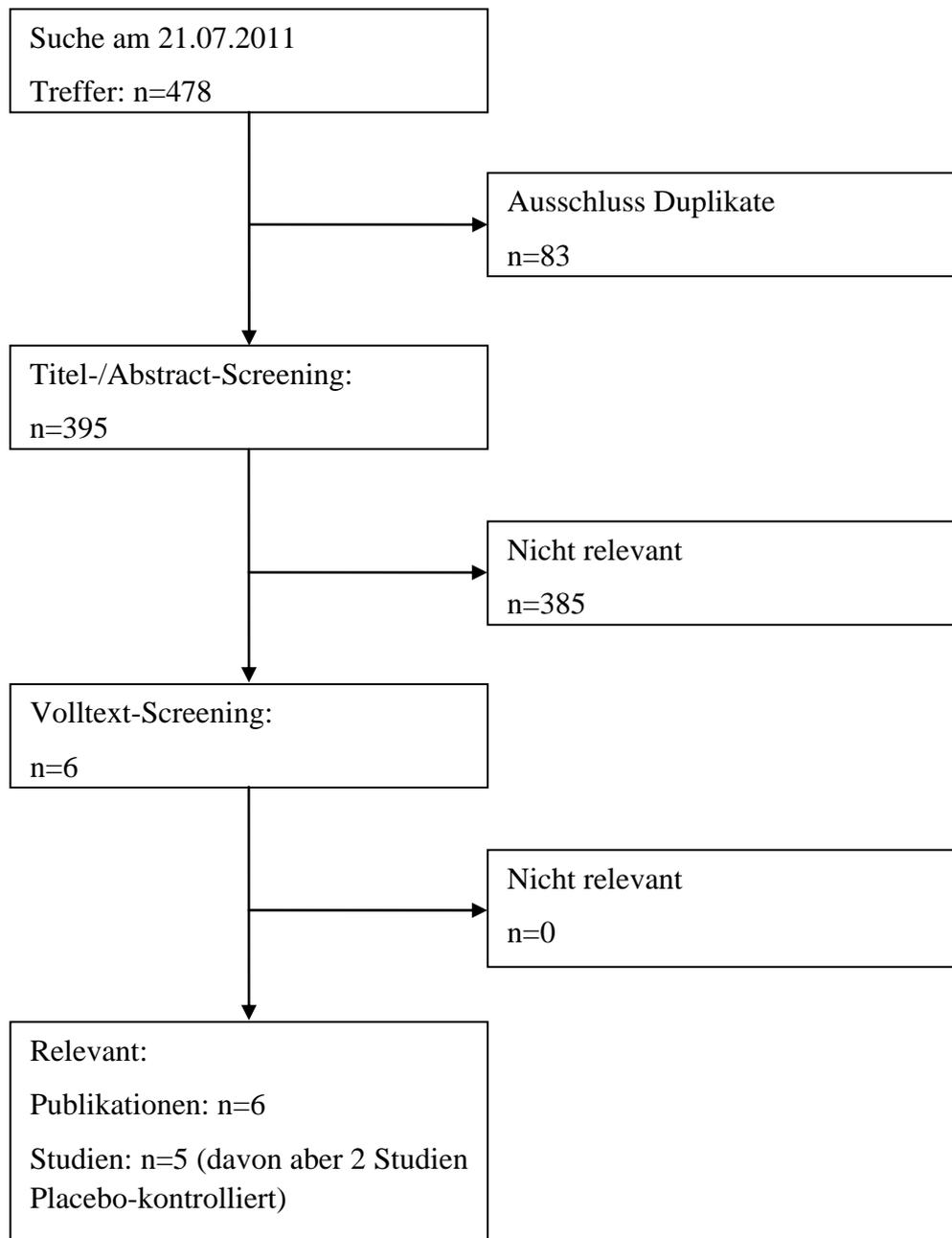
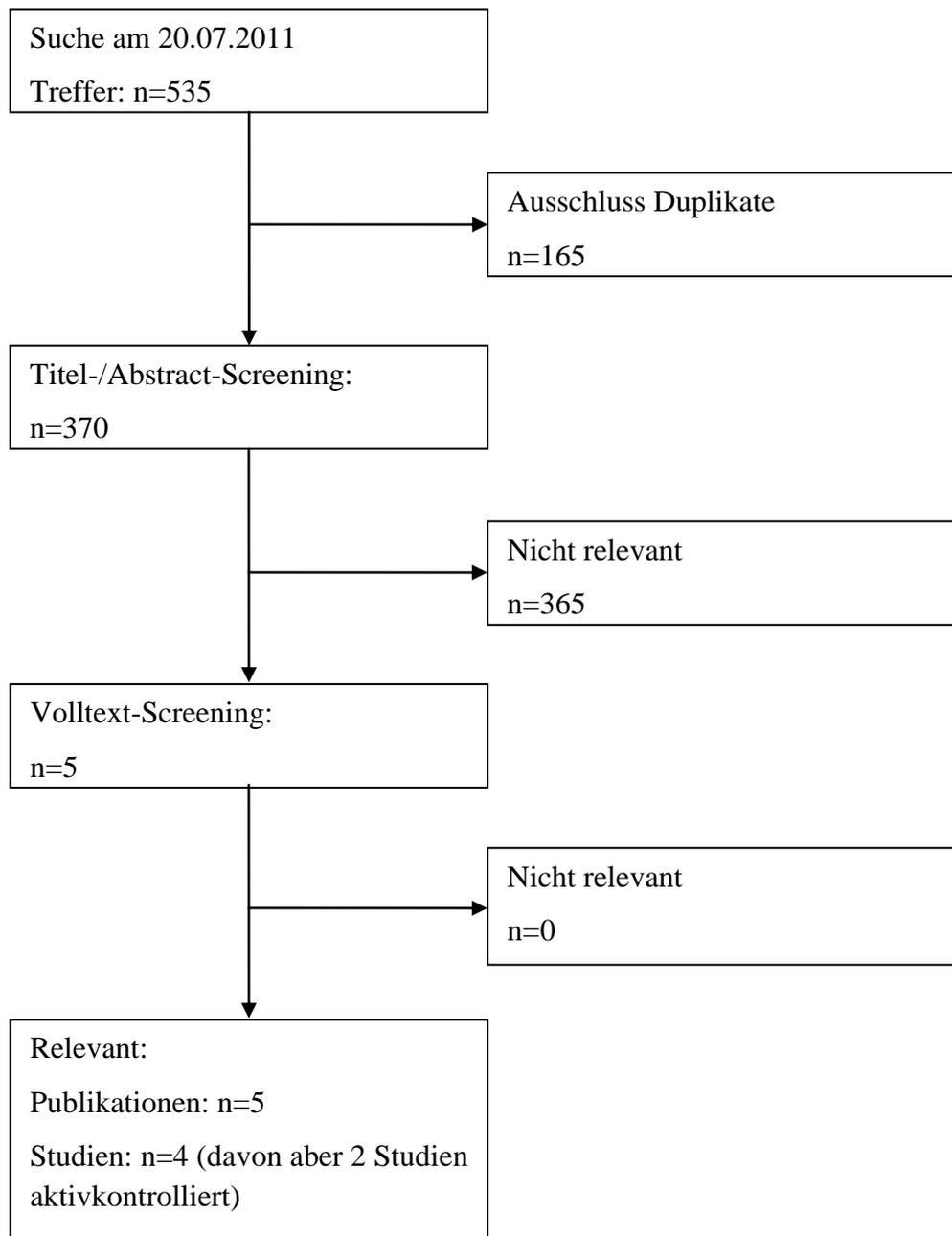


Abbildung 4-25: Flussdiagramm der bibliografischen Literaturrecherche zu Glatirameracetat-Studien (Placebo-kontrolliert)



4.3.2.1.1.3 Studien aus der Suche in Studienregistern

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-4) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde (sofern eine solche durchgeführt 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 4-51: Relevante Studien aus der Suche in Studienregistern – Studien für indirekte Vergleiche

Studienbezeichnung	Identifikationsorte (Name der Studienregister und Angabe der Zitate^a)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d.^b)
FTY720D2301 FREEDOMS	clinicaltrials.gov [NCT00289978], Registereintrag und Ergebnisse (25); International Clinical Trials Registry Platform [NCT00289978], Registereintrag (26)	ja	ja
FTY720D2302 TRANSFORMS	clinicaltrials.gov [NCT00340834], Registereintrag und Ergebnisse (27); International Clinical Trials Registry Platform [NCT00340834], Registereintrag (28)	ja	ja
EVIDENCE	clinicaltrials.gov [NCT00292266], Registereintrag (45); International Clinical Trials Registry Platform [NCT00292266], Registereintrag (46)	nein	ja
ABOVE	clinicaltrials.gov [NCT00206648], Registereintrag (47); International Clinical Trials Registry Platform [NCT00206648], Registereintrag (48); clinicalstudyresults.org [NCT00206648] Synopse (49)	nein	nein
Johnson et al., 1995	clinicaltrials.gov [NCT00004814], Registereintrag (50); International Clinical Trials Registry Platform [NCT00004814], Registereintrag (51)	nein	ja
BEYOND	clinicaltrials.gov [NCT00099502], Registereintrag (52); International Clinical Trials Registry Platform [NCT00099502], Registereintrag (53);	nein	ja

Studienbezeichnung	Identifikationsorte (Name der Studienregister und Angabe der Zitate ^a)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d. ^b)
	clinicalstudyresults.org [NCT00099502] Synopse (54)		
BECOME	clinicaltrials.gov [NCT00176592], Registereintrag (55); International Clinical Trials Registry Platform [NCT00176592], Registereintrag (56)	nein	ja
REGARD	clinicaltrials.gov [NCT00078338], Registereintrag (57) International Clinical Trials Registry Platform [NCT00078338], Registereintrag (58)	nein	ja
k.A.	clinicaltrials.gov [NCT00202995], Registereintrag (59); International Clinical Trials Registry Platform [NCT00202995], Registereintrag (60)	nein	nein
k.A	clinicaltrials.gov [NCT00101959], Registereintrag (61)	nein	nein
<p>a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. b: n. d. = bibliografische Literaturrecherche nicht durchgeführt.</p>			

In Studienregistern wurden folgende weitere abgeschlossene RCT identifiziert:

NCT00202995

NCT00101959

NCT00206648 (ABOVE)

Eine Einbeziehung in die Bewertung der Studien NCT00202995 und NCT00101959 ist nicht möglich, da keine Ergebnisdaten zu diesen Studien verfügbar sind. Die Studie NCT00206648 (ABOVE) wurde nicht eingeschlossen, da diese wegen fehlender wissenschaftlicher Integrität abgebrochen wurde.

4.3.2.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

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

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Datenquellen. 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 4-52: Studienpool – Studien für indirekte Vergleiche

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
Placebo-kontrolliert						
FREEDOMS	ja	ja	nein	ja [Studienbericht und Zusatzauswertungen] (22;29)	ja [NCT00289978] (25;26)	ja [Kappos et al., 2010](30)
MSCRG	nein	nein	ja	nein	nein	ja [Jacobs et al., 1996, Miller et al., 2011, Rudick et al., 1997, 2001, 2010; Jacobs et al., 1995] (41;62-66)
Johnson et al. (1995)	nein	nein	ja	nein	ja [NCT00004814] (50;51)	ja [Johnson et al., 1995, 2001; Liu et al., 2000] (67-69)
The IFNB MS Study Group	nein	nein	ja	nein	nein	ja [The IFNB Multiple Sclerosis Study Group, 1993; Anonymous, 1995, Duquette et al., 1993, 2001] (39;70-72)
PRISMS	nein	nein	ja	nein	nein	ja [The PRISMS Study Group, Ebers et al., 1998; Liu et al., 1999] (40;73)
Bornstei	nein	nein	ja	nein	nein	ja [Bornstein

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
n et al. (1987)						et al., 1987] (74)
aktiv-kontrolliert, zweckmäßige Vergleichstherapie(n)						
TRANSFORMS	ja	ja	nein	ja [Studienbericht. und Zusatzauswertungen] (23;29)	ja [NCT00340834] (27;28)	ja [Cohen et al., 2010](31)
EVIDENCE	nein	nein	ja	nein	ja [NCT00292266] (45;46)	ja [Panitch et al., 2002; Schwid und Panitch, 2007; Schwid et al., 2005; Sandberg-Wollheim et al., 2005; Panitch et al., 2005] (75-79)
INCOMIN	nein	nein	ja	nein	nein	ja [Durelli et al., 2002] (43)
BEYOND	nein	nein	ja	nein	ja [NCT00099502] (52;53;54)	ja [O'Connor et al., 2009] (80)
BECOME	nein	nein	ja	nein	ja [NCT00176592] (55;56)	ja [Cadavid et al., 2009] (81)
REGARD	nein	nein	ja	nein	ja [NCT00078338] (57;58)	ja [Mikol et al., 2008] (82)
Etemadifar et al. (2006)	nein	nein	ja	nein	nein	ja [Etemadifar et al., 2006] (83)
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.</p> <p>b: Studie, für die der Unternehmer Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p>						

4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

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

4.3.2.1.2.1 Studiendesign und Studienpopulationen

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

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

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

Tabelle 4-53: Charakterisierung der eingeschlossenen Studien – Studien für indirekte Vergleiche

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
Placebo-kontrolliert						
FREEDOMS	RCT, doppelblind, parallel	Erwachsene mit RRMS 1 Schub im letzten Jahr, oder 2 Schübe in den letzten 2 Jahren EDSS 0 - 5,5	Fingolimod 1,25 mg (n=429) Fingolimod 0,5 mg (n=425) Placebo (n=418)	Screening: 1 Monat Behandlung: 24 Monate	Weltweit in 22 Ländern 1/2006 – 7/2009	Schubhäufigkeit; Behinderungspro- gression, MRT-Para- meter, Kognition, uner- wünschte Ereignisse
MSCRG	RCT, doppelblind, Placebo-kontrolliert	Erwachsene (18- 55 J.) mit RRMS Mind. 2 Schübe in den letzten 3 Jahren EDSS 1,0 - 3,5	IFN- β -1a 30 μ g i. m. (n=158) Placebo (n=143)	Behandlung: 24 Monate	USA: 4 Zentren 11/1990 – 1995	Dauer bis zur EDSS- Verschlechterung um mindestens 1,0;
Johnson et al. (1995)	RCT, doppelblind, Placebo-kontrolliert	Erwachsene (18- 45 J.) mit RRMS Mind. 2 Schübe in den letzten 2 Jahren EDSS 0 - 5,0	Glatirameracetat 20 mg s. c. (n=125) Placebo (n=126)	Behandlung: 24 Monate	USA: 11 Zentren 10/1991 – 1994	Schubhäufigkeit; Anteil der Patienten ohne Schübe, Zeit bis 1. Schub;
The IFNB MS Study Group	RCT, doppelblind, parallel, Placebo- kontrolliert Gepoolte Ergebnisse zweier identischer Studien	Erwachsene (18- 50 J.) mit RRMS Mind. 2 Schübe in den letzten 2 Jahren EDSS \leq 5,5	IFN- β -1b 50 μ g s. c. (n=125) IFN- β -1b 250 μ g s. c. (n=124) Placebo (n=123)	Behandlung: 2 Jahre Danach Weiterbehandlung möglich bis 5,5 Jahre	USA und Canada: 11 Zentren 06 / 1988 – 1995	Jährliche Schubrate, Anteil schubfreier Patienten; Zeitraum (Tage) bis zum 1. Schub
PRISMS	RCT, doppelblind,	Erwachsene mit	IFN- β -1a 22 μ g s. c.	Behandlung:	Weltweit in	Schubhäufigkeit;

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

	Placebo-kontrolliert	RRMS; Mind. 2 Schübe in den letzten 2 Jahren EDSS 0 - 5,0	(n=189) IFN-β-1a 44 µg s. c. (n=184) Placebo (n=187)	24 Monate	9 Ländern 05/1994 – 1997	Verschlechterung der Behinderungspro- gression;
Bornstein et al. (1987)	RCT, doppelblind, Placebo-kontrolliert	Erwachsene (20- 35 J.) mit RRMS Mindestens 2 Schübe in den letzten 2 Jahren EDSS ≤ 6,0	Glatirameracetat 20 mg s. c. (n=25) Placebo (n=25)	Behandlung: 24 Monate	USA: 1 Zentrum	Schubhäufigkeit; Behinderungsprogressi on
Aktiv-kontrolliert						
TRANSFORMS	RCT, doppelblind, parallel	Erwachsene mit RRMS 1 Schub im letzten Jahr, oder 2 Schübe in den letzten 2 Jahren EDSS 0 - 5,5	Fingolimod 1,25 mg (n=426) Fingolimod 0,5 mg (n=431) IFN-β-1a 30 µg i. m. (n=435) Und jeweiliges Placebo (Double- Dummy)	Screening: 1 Monat Behandlung: 12 Monate	Weltweit in 18 Ländern 5/2006 – 11/2008	Schubhäufigkeit; Behinderungspro- gression, MRT-Para- meter, Kognition, uner- wünschte Ereignisse
EVIDENCE	RCT, auswerterver- blinded, parallel,	IFN-naive Patienten (18-55 J.) mit RRMS Mind. 2 Schübe in den letzten 2 Jahren EDSS 0 - 5,5	IFN-β-1a 44 µg s. c. (n=339) IFN-β-1a 30 µg i. m. (n=338)	Screening: 1 Monat Behandlung: 24 Wochen	Europa: 15 Zentren, Kanada: 5 Zentren, USA: 36 Zentren	Anteil der Patienten frei von Schüben über 24 Wochen (OR); Schubhäufigkeit und – schwere;
INCOMIN	RCT, parallel, prospektiv	Erwachsene (18- 50 J.) mit RRMS Mindestens 2 Schübe in den letzten 2 Jahren Mindestens 30 Tage vor Studieneintritt	IFN-β-1a i. m. 30 µg i. m. (n=92) IFN-β-1b 250 µg s. c. (n=96)	Behandlung: 24 Monate	Italien: 15 Zentren 10/1997 – 06/2001	Anteil der Patienten frei von Schüben und von neuen T2-Läsionen; Jährliche (behandelte) Schubrate; Anteil der Patienten ohne fortschreitende

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

		kein Schub EDSS 1,0 – 3,5				Behinderung; Dauer bis zum Fortschreiten der Behinderung
BEYOND	RCT, doppelblind für Interferon- Dosierung, einfachblind für Bewertung, parallel	Erwachsene (18- 55 J.) mit RRMS Treatment-naiv Mind. 1 Schub im letzten Jahr EDSS 0 – 5,0	IFN-β-1b 500 µg s. c. (n=899) IFN-β-1b 250 µg s. c. (n=897) Glatirameracetat 20 mg s. c. (n=448)	Behandlung: 2 Jahre	Weltweit in 26 Ländern 11/2003 – 08/2007	Rückfallrisiko; EDSS-Veränderung; MRT-Parameter
BECOME	RCT, offen, einfachblind für Bewertung, parallel	Erwachsene (18- 55 J.) mit RRMS oder CIS für MS Behandlungs-naiv	IFN-β-1b 250 µg s. c. (n=36) Glatirameracetat 20 mg s. c. (n=39)	Behandlung: 1 Jahr	USA 01/2003-01/2006	Anzahl kombinierte aktive Läsionen; Neue Läsionen und Schübe;
REGARD	RCT, offen, parallel	Erwachsene (18- 60 J.) mit RRMS IFN-β-1a und Glatirameracetat-naiv Mind. 1 Schub im letzten Jahr EDSS 0 - 5,5	IFN-β -1a 44 µg s. c. (n = 386) Glatirameracetat 20 mg s. c. (n=378)	Screening: 4 Wochen Behandlung: 96 Wochen	Weltweit in 14 Ländern 02/2004 – 09/2007	Dauer bis zum ersten Schub; Anzahl neuer verstärkter T2-Läsionen und Volumen- Veränderungen;
Etemadifar et al. (2006)	RCT, einfachblind, parallel	Patienten (15-50 J.) mit RRMS Mindestens 2 Schübe in den letzten 2 Jahren EDSS ≤ 5	IFN-β-1b 250 µg s. c. (n=30) IFN-β-1a 30 µg i. m. (n=30) IFN-β-1a 44 µg s. c. (n=30)	Behandlung: 24 Monate	Iran Ab 9/2002	Schubhäufigkeit; EDSS-Veränderung;

Tabelle 4-54: Charakterisierung der Interventionen – Studien für indirekte Vergleiche

Studie	Gruppe 1	Gruppe 2	Gruppe 3	Vorbehandlung der MS mit zugelassener Therapie n (%)
Placebo-kontrolliert				
FREEDOMS	Fingolimod 0,5 mg 1x tgl.	Placebo 1x tgl.	Fingolimod ^a 1,25 mg 1x tgl.	520 (40,9)
MSCRG	IFN-β-1a 30 µg i. m. 1 x / Woche	Placebo 1 x / Woche	-	k.A. ^d
Johnson et al. (1995)	Glatirameracetat 20 mg s. c. 1x tgl.	Placebo 1x tgl.	-	k.A. ^d
The IFNB MS Study Group	IFN-β-1b 50 µg s. c. alle 2 Tage	IFN-β-1b 250 µg s. c. alle 2 Tage	Placebo alle 2 Tage	0 (0,0) ^c
PRISMS	IFN-β-1a 22 µg s. c. 3 x / Woche	IFN-β-1a 44 µg s. c. 3 x / Woche	Placebo s. c. 3 x / Woche	(3%)
Bornstein et al. (1987)	Glatirameracetat 20 mg s. c. 1x tgl.	Placebo 1x tgl.	-	k.A. ^d
Aktiv-kontrolliert				
TRANSFORMS	Fingolimod 0,5 mg 1x tgl. + Placebo i. m. 1 x / Woche	IFN-β-1a 30 µg i. m. 1x/Woche + Placebo 1x tgl.	Fingolimod ^a 1,25 mg 1x tgl. + Placebo i. m. 1 x / Woche	732 (56,7)
EVIDENCE	IFN-β-1a 44 µg s. c. 3 x / Woche	IFN-β-1a 30 µg i. m. 1 x / Woche	-	k.A. ^b
INCOMIN	IFN-β-1a 30 µg i. m. 1 x / Woche	IFN-β-1b 250 µg s. c. alle 2 Tage	-	0 (0%) ^c
BEYOND	IFN-β-1b 500 µg s. c. Alle 2 Tage	IFN-β-1b 250 µg s. c. Alle 2 Tage	Glatirameracetat 20 mg s. c. 1x tgl.	0 (0%) ^c
BECOME	IFN-β-1b 250 µg s. c. Alle 2 Tage	Glatirameracetat 20 mg s. c. 1x tgl.	-	0 (0%) ^c
REGARD	IFN-β-1a 44 µg s. c. 3 x / Woche	Glatirameracetat 20 mg s. c. 1x tgl.	-	0 (0%) ^c
Etemadifar et al. (2006)	IFN-β-1b 250 µg s. c. alle 2 Tage	IFN-β-1a 30 µg i. m. 1 x / Woche	IFN-β-1a 44 µg s. c. 3 x / Woche	k.A. ^d

- a: Ergebnisse der Gruppe 3 bei FREEDOMS und TRANSFORMS werden nicht weiter dargestellt, da die untersuchte Dosierung nicht zur Zulassung eingereicht wurde
- b: IFN- β -Vorbehandlung war ausgeschlossen
- c: immunmodulatorische Vorbehandlung ausgeschlossen
- d: aus der Publikation war keine Angabe ersichtlich

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

Tabelle 4-55: Charakterisierung der Studienpopulationen – Studien für indirekte Vergleiche

Studie Gruppe	N	Alter (Jahre) (M ± SD)	Geschlecht w / m (%)	Dauer Erkrankung (Jahre) (M ± SD)	Baseline EDSS (M ± SD)	Anzahl Schübe in den letzten 2 Jahren (M ± SD)	Teilnehmer ohne Gd-anreichernde T1 Läsionen n (%)	Teilnehmer ohne Vorbehandlung n (%)
Placebo-kontrolliert								
FREEDOMS	1272	37,1 ± 8,76	69,9 / 30,1	8,2 ± 6,60 ^a	2,40 ± 1,32	2,1 ± 1,19	782/1264 (61,9)	752 (59,1)
(Fingolimod 1,25 mg) ^b	429	37,4 ± 8,91	68,8 / 31,2	8,4 ± 6,86	2,41 ± 1,36	2,1 ± 1,25	257/424 (60,6)	259 (60,4)
Fingolimod 0,5 mg	425	36,6 ± 8,77	69,6 / 30,4	8,0 ± 6,60	2,30 ± 1,29	2,1 ± 1,13	263/424 (62,0)	244 (57,4)
Placebo	418	37,2 ± 8,60	71,3 / 28,7	8,1 ± 6,35	2,49 ± 1,29	2,2 ± 1,19	262/416 (63,0)	249 (59,6)
MSCRG	301						k.A.	k.A.
IFN-β-1a 30 µg i. m.	158	36,7 ± 0,57 ^f	74,6 / 25,4	6,6 ± 0,49 ^f	2,4 ± 0,06 ^f	1,2 ± 0,05 ^{c, f}		
Placebo	143	36,9 ± 0,64 ^f	72,0 / 28,0	6,4 ± 0,46 ^f	2,3 ± 0,07 ^f	1,2 ± 0,05 ^{c, f}		
Johnson et al. (1995)	251						k.A.	k.A.
Glatirameracetat 20 mg s. c.	125	34,6 ± 6,0	70,4 / 29,6	7,3 ± 4,9	2,8 ± 1,2	2,9 ± 1,3		
Placebo	126	34,3 ± 6,5	76,2 / 23,8	6,6 ± 5,1	2,4 ± 1,3	2,9 ± 1,1		
The IFNB MS Study Group.	372		69,6 / 30,4				k.A.	372 (100,0)
IFN-β-1b 50 µg s. c.	125	35,3 ± 0,7 ^f	68,0 / 32,0	4,7 ± 0,4 ^f	2,9 ± 0,1 ^f	3,3 ± 0,1 ^f	k.A.	125 (100,0)
IFN-β-1b 250 µg s. c.	124	35,3 ± 0,6 ^f	69,4 / 30,6	4,7 ± 0,4 ^f	3,0 ± 0,2 ^f	3,4 ± 0,2 ^f	k.A.	124 (100,0)
Placebo	123	36,0 ± 0,6 ^f	71,5 / 28,5	3,9 ± 0,3 ^f	2,8 ± 0,1 ^f	3,6 ± 0,1 ^f	k.A.	123 (100,0)

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

PRISMS	560	34,9e (29,1-40,0) ^d	69 / 31	5,3 e(2,8-10) ^d	2,5 ± 1,2	3,0 ± 1,2	k.A.	n.d. (97,0)
IFN-β-1a 22 µg s. c.	189	34,8 e(29,3-39,8) ^d	67 / 33	5,4 e(3-11,2) ^d	2,5 ± 1,2	3,0 ± 1,1		
IFN-β-1a 44 µg s. c.	184	35,6 e(28,4-41,0) ^d	66 / 34	6,4 e(2,9-10,03) ^d	2,5 ± 1,3	3,0 ± 1,1		
Placebo	187	34,6 e(28,8-40,4) ^d	75 / 25	4,3 e(2,4-8,4) ^d	2,4 ± 1,2	3,0 ± 1,3		
Bornstein et al. (1987)	50		58,0 / 42,0				k.A.	k.A.
Glatiramer- acetat 20 mg s. c.	25	30,0	56,0 / 44,0	4,9	2,9	3,8		
Placebo	25	31,0	60,0 / 40,0	6,1	3,2	3,9		
Aktiv-kontrolliert								
TRANSFORMS	1292	36,2 ± 8,50	67,3 / 32,7	7,4 ± 6,16 ^a	2,21 ± 1,30	2,2 ± 1,61	826/1264 (65,3)	560 (43,3)
(Fingolimod 1,25 mg) ^b	426	35,8 ± 8,39	68,8 / 31,2	7,3 ± 5,96	2,21 ± 1,31	2,2 ± 1,19	270/412 (65,5)	177 (41,5)
Fingolimod 0,5 mg	431	36,7 ± 8,81	65,4 / 34,6	7,5 ± 6,20	2,24 ± 1,33	2,3 ± 2,20	288/427 (67,4)	193 (44,8)
IFN-β-1a 30 µg i. m.	435	36,0 ± 8,29	67,8 / 32,2	7,4 ± 6,34	2,19 ± 1,26	2,3 ± 1,22	268/425 (63,1)	190 (43,7)
EVIDENCE	677						k.A.	k.A.
IFN-β-1a 44 µg s. c.	339	38,3	74,9 / 25,1	6,5	2,3	2,6		
IFN-β-1a 30 µg i. m.	338	37,4	74,6 / 25,4	6,7	2,3	2,6		
INCOMIN	188						k.A.	k.A.
IFN-β-1a 30 µg i. m.	92	34,9 ± 7,9	62,0 / 38,0	6,7 ± 5,4		1,38 ± 0,52		
IFN-β-1b 250 µg s. c.	96	38,8 ± 7,1	68,8 / 31,2	5,9 ± 4,2		1,52 ± 0,67		

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

BEYOND	2244						1224/ 2244 (54,5)	2244 (100,0)
IFN-β-1b 500 µg s. c.	899	35,9	70,0 / 30,0	5,4 (1-8) ^d	2,33 (1,5-3) ^d	1,6 (1-2) ^c	494 / 899 (54,9)	899 (100,0)
IFN-β-1b 250 µg s. c.	897	35,8	69,9 / 30,1	5,3 (1-7) ^d	2,35 (1,5-3) ^d	1,6 (1-2) ^c	484 / 897 (54,0)	897 (100,0)
Glatiramer- acetat 20 mg s. c.	448	35,2	68,3 / 31,7	5,1 (1-7) ^d	2,28 (1,5-3) ^d	1,6 (1-2) ^c	246 / 448 (54,9)	448 (100,0)
BECOME	75 / 64 [§]						k.A.	74 (98,7)
IFN-β-1b 250 µg s. c.	36 / 29 [§]	36,0	75,0	0,90 ^e	2,0 ^e	2,72		
Glatiramer- acetat 20 mg s. c.	39 / 35 [§]	36,0	64,1	1,20 ^e	2,0 ^e	2,87		
REGARD	764		70,5 / 29,5				k.A.	764 (100,0)
IFN-β-1a 44 µg s. c.	386	36,7 ± 9,8	69,2 / 30,8	5,93 ± 6,25	2,35 ± 1,28	0,97		386 (100,0)
Glatiramer- acetat 20 mg s. c.	378	36,8 ± 9,5	72,0 / 28,0	6,55 ± 7,10	2,33 ± 1,31	1,01		378 (100,0)
Etemadifar et al (2006)	90	28,5 ± 7,0	75,6 / 24,4				k.A.	k.A.
IFN-β-1b 250 µg s. c.	30	29,9 ± 1,4	70,0 / 30,0	3,7 ± 2,3	1,9 ± 0,7	2,2 ± 0,7		
IFN-β-1a 30 µg i. m.	30	28,1 ± 1,2	80,0 / 20,0	2,9 ± 2,3	1,9 ± 1,1	2,0 ± 0,8		
IFN-β-1a 44 µg s. c.	30	27,4 ± 1,2	76,7 / 23,3	3,0 ± 2,2	2,1 ± 1,0	2,4 ± 1,0 ^c		

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

- a: Seit Diagnosestellung
- b: Ergebnisse der Gruppe 3 werden nicht weiter dargestellt
- c: innerhalb eines Jahres
- d: IQR-Werte (Interquartilsabstand)
- e: Median
- f: SEM Standardfehler des Mittelwertes
- g: 2 Jahre
- k.A.: keine Angabe

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

Alle Studien, die in den indirekten Vergleich eingeschlossen wurden, waren RCT zu Fingolimod, IFN- β -Präparaten und Glatirameracetat. Die Studien FREEDOMS, MSCRG, Johnson et al., The IFNB MS Study Group, PRISMS und Bornstein et al. waren Placebo-kontrolliert, davon testeten Johnson et al. und Bornstein et al. Glatirameracetat. Die aktiv-kontrollierten Studien waren TRANSFORMS, EVIDENCE, INCOMIN, BEYOND, BECOME, REGARD und Etemadifar, in den Studien BEYOND, BECOME und REGARD waren Arme mit Glatirameracetat enthalten. Die Studiendauer lag bei 12 bzw. 24 Monaten. Die Studienpopulationen umfassten RRMS Patienten.

Die Studien unterscheiden sich neben den Behandlungsarmen vor allem in den Einschlusskriterien hinsichtlich der Vorbehandlung. EVIDENCE, FREEDOMS; MSCRG, Johnson et al., The IFNB MS Study Group, TRANSFORMS, Etemadifar et al. und Bornstein et al. erlaubten eine Vorbehandlung mit Immunmodulatoren, EVIDENCE allerdings keine Vorbehandlung mit IFN- β . Meist musste die Vorbehandlung aber einige Monate vor Studienbeginn beendet werden. Die Zeiträume wurden dabei unterschiedlich gewählt. Details können dem Bericht zum indirekten Vergleich entnommen werden (RTI Bericht 2011) (24).

Ein weiterer Unterschied ergibt sich in der Studienpopulation aus den unterschiedlichen, zur jeweiligen Studienzeit gültigen MS-Diagnosekriterien. In Studien, die vor 2001 rekrutierten, fanden noch die Poser-Kriterien Anwendung, in späteren Studien die McDonald-Kriterien. Daher können sich die Populationen hinsichtlich Krankheitsschwere unterscheiden (84).

4.3.2.1.2.2 Verzerrungspotenzial auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-56: Verzerrungspotenzial auf Studienebene – Studien für indirekte Vergleiche

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial auf Studienebene
			Patient	Behandler			
Placebo-kontrolliert							
FREEDOMS	ja	ja	ja	ja	nein	nein	niedrig
MSCRG	ja	unklar	ja	ja	nein	nein	niedrig
Johnson et al. (1995)	ja	ja	ja	ja	nein	nein	niedrig
The IFNB MS Study Group	ja	ja	ja	ja	nein	nein	niedrig
PRISMS	ja	ja	ja	ja	nein	nein	niedrig
Bornstein et al. (1987)	unklar	unklar	ja	ja	nein	nein	niedrig
Aktiv-kontrolliert							
TRANSFORMS	ja	ja	ja	ja	nein	nein	niedrig
EVIDENCE	ja	ja	nein	nein	nein	nein	niedrig
INCOMIN	ja	ja	nein	nein	nein	nein	hoch
BEYOND	ja	ja	ja	ja	nein	nein	niedrig
BECOME	unklar	unklar	nein	nein	nein	nein	niedrig
REGARD	ja	unklar	nein	nein	nein	nein	niedrig
Etemadifar et al. (2006)	unklar	unklar	nein	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial auf Studienebene sowohl für FREEDOMS als auch TRANSFORMS als „niedrig“ eingestuft, da jeweils eine adäquate Erzeugung der Randomisierungssequenz vorlag, die Gruppenzuteilung verdeckt erfolgte, sowohl Patient als auch Behandler verblindet waren (doppelblind), für die unterschiedlichen Interventionen in der TRANSFORMS ein Double-Dummy Design gewählt wurde und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte gab.

Für die Studien Johnson et al. (1995), IFNB MS Study Group, PRISMS und BEYOND gilt, soweit das aus den Publikationen erkennbar ist, ebenfalls, dass das Verzerrungspotenzial auf Studienebene als „niedrig“ eingestuft wird, da jeweils eine adäquate Erzeugung der Randomisierungssequenz vorlag, die Gruppenzuteilung verdeckt erfolgte, sowohl Patient als auch Behandler verblindet waren (doppelblind) und es weder Hinweise auf ergebnisgesteuerte

Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte gab. Diese Punkte gelten entsprechend auch für die Studien MSCRG und Bornstein et al. (1987). Lediglich die Erzeugung der Randomisierungssequenz bzw. die Verdeckung der Gruppenzuteilung war teilweise nicht klar aus der Publikation erkennbar. Der Einfluss auf das Ergebnis wird als irrelevant betrachtet.

Für die Studien EVIDENCE, BECOME, REGARD und Etemadifar et al. (2006) gilt ebenfalls die Einstufung des Verzerrungspotentials als niedrig. Zwar waren Patienten und Behandler nicht verblindet (bei Etemadifar nur Patienten nicht verblindet), aber die Endpunkterheber waren verblindet. Ob die Erzeugung der Randomisierungssequenz adäquat war und die Gruppenzuteilung verdeckt erfolgte war bei diesen Studien teilweise unklar, der Einfluss auf das Ergebnis wird als irrelevant betrachtet. Ansonsten gab es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte.

Das Verzerrungspotential der Studie INCOMIN wurde als hoch eingestuft, da alle Endpunkte unverblindet erhoben wurden.

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 4-57: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	Jährliche Schubrate	Schubfreie Patienten	Behinderungsprogression ^a	Behandlungsabbrüche wg UE
Placebo-kontrolliert				
FREEDOMS	ja	ja	ja	ja
MSCRG	ja	ja	nein	ja
Johnson et al. (1995)	ja	ja	ja	nein
The IFNB MS Study Group	ja	ja	ja	ja
PRISMS	ja	ja	nein	ja
Bornstein et al. (1987)	ja	ja	nein	nein
Aktiv-kontrolliert				
TRANSFORMS	ja	ja	Ja	ja
EVIDENCE	ja	ja	nein	ja
INCOMIN	ja	ja	nein	ja
BEYOND	ja	ja	ja	ja
BECOME	ja	ja	nein	nein
REGARD	ja	ja	nein	ja
Etemadifar et al. (2006)	nein	ja	nein	nein
a: nach 3 Monaten bestätigte Behinderungsprogression				

4.3.2.1.3.1 Endpunkte – indirekte Vergleiche aus RCT

4.3.2.1.3.1.1 Jährliche Schubrate– indirekte Vergleiche aus RCT

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

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

Tabelle 4-58: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Referenzen Studien	Intervention	IFN- β -1a i. m.	IFN- β -1a s. c.	IFN- β -1b s. c.	Glatiramer acetat	Placebo
Placebo-kontrolliert							
1	FREEDOMS	•					•
1	MSCRG		•				•
2	Johnson et al. (1995) Bornstein et al. (1987)					• •	• •
1	The IFNB MS Study Group				•		•
1	PRISMS			•			•
Aktiv-kontrolliert							
1	TRANSFORMS	•	•				
1	EVIDENCE		•	•			
1	INCOMIN		•		•		
2	BEYOND BECOME				• •	• •	
1	REGARD			•		•	

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

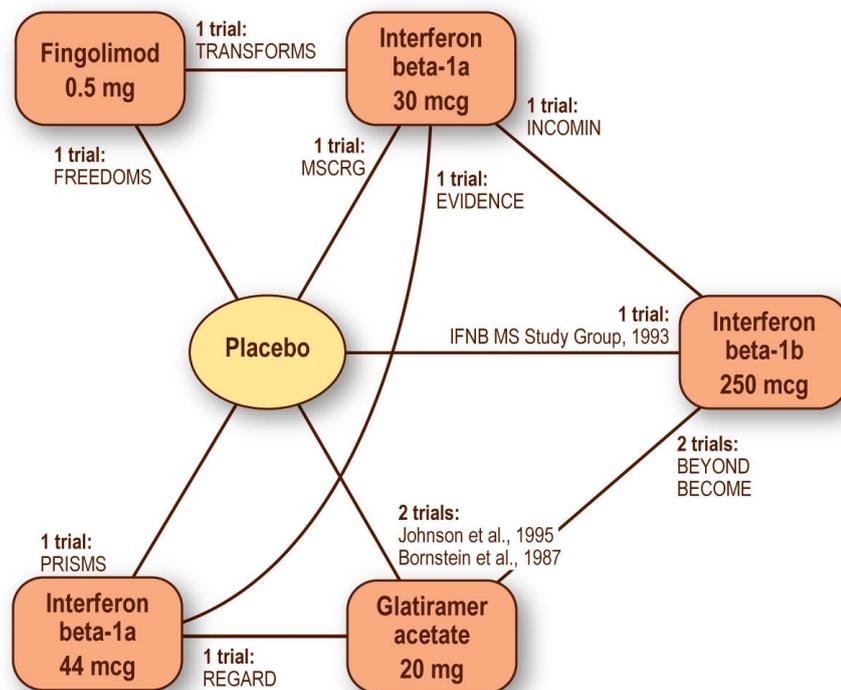


Abbildung 4-26: Netzwerkstruktur für Endpunkt „Jährliche Schubrate“

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

Tabelle 4-59: Operationalisierung von „Jährliche Schubrate“

Studie	Operationalisierung
Placebo-kontrolliert	
FREEDOMS	<p>Aggregierte jährliche Schubrate: Gesamtzahl aller bestätigten Schübe (aus CSR) aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25</p> <p>Schubdefinition Ein Schub ist allgemein definiert als neue, verschlechterte oder erneut auftretende neurologische Symptome, die mindestens 30 Tage nach Beginn des vorausgegangenen Schubs auftraten, mindestens 24 Stunden anhielten, ohne Fieber oder Infektion verliefen. Ein bestätigter Schub liegt vor, wenn die allgemeinen Symptome von einem Anstieg um $\geq 0,5$ EDSS-Punkte oder um 1 Punkt in zwei Funktionssystem-Werten oder um 2 Punkte in einem Funktionssystem-Wert (mit Ausnahme von Darm/Blase und psychischen Systemen) begleitet waren.</p>
MSCRG	<p>Aggregierte jährliche Schubrate: Gesamtzahl aller bestätigten Schübe aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25</p> <p>Gesamtzahl aller Schübe aller Patienten Mittels Imputation bestimmt anhand der Personenjahre und der nicht-adjustierten Schubrate</p> <p>Gesamtzahl der Studientage aller Patienten Mittels Imputation bestimmt anhand der Tabelle zu Wochen der Nachbeobachtung (Dropouts erhielten die Hälfte der Zeit bis zu dem Zeitpunkt)</p> <p>Schubdefinition: Schübe im Studienverlauf waren definiert als neue neurologische Symptome oder Verschlechterung bestehender neurologischer Symptome, die über mindestens 48 Stunden anhielt(en), bei einem Patienten, der in den vorangegangenen 30 Tagen neurologisch stabil war oder eine Besserung gezeigt hatte, begleitet von einer objektiven Änderung bei der neurologischen Untersuchung (Verschlechterung auf der EDSS-Skala um 0,5 Punkte oder Verschlechterung des Punktwerts für das pyramidale, zerebellare, den Hirnstamm betreffende oder visuelle funktionelle System um mindestens 1,0 Punkte).</p>
Johnson et al. (1995)	<p>Aggregierte jährliche Schubrate: Gesamtzahl aller bestätigten Schübe aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25</p> <p>Gesamtzahl aller Schübe aller Patienten In der Veröffentlichung angegeben</p> <p>Gesamtzahl der Studientage aller Patienten Mittels Imputation bestimmt anhand der Anzahl der Schübe und der adjustierten (nach Baseline-Kovariablen) jährlichen Schubrate (ARR)</p> <p>Schubdefinition: Ein Schub war definiert als Auftreten oder erneutes Auftreten einer oder mehrerer neurologischer Störungen, die über mindestens 48 Stunden anhielten und denen unmittelbar eine Phase der relativen Stabilität oder Verbesserung des neurologischen Zustands von einer Dauer von mindestens 30 Tagen vorausgegangen war. Zur Bestätigung eines Schubs wurde eine objektive</p>

	Veränderung bei der neurologischen Untersuchung verwendet, die einem Anstieg um $\geq 0,5$ EDSS-Punkten, 2 Punkten bei einem der sieben funktionellen Systeme oder 1 Punkt bei zwei oder mehr funktionellen Systemen entsprach.
IFNB MS Study Group	<p>Aggregierte jährliche Schubrate:</p> <p>Gesamtzahl aller bestätigten Schübe aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25</p> <p>Gesamtzahl aller Schübe aller Patienten In der Veröffentlichung angegeben</p> <p>Gesamtzahl der Studientage aller Patienten In der Veröffentlichung angegeben</p> <p>Schubdefinition:</p> <p>Ein Schub war definiert als neues Symptom oder Verschlechterung eines alten Symptoms, das/die auf die MS zurückführbar war, von einer dazu passenden neuen neurologischen Störung begleitet war, über einen Zeitraum von mindestens 24 Stunden anhielt, ohne Fieber verlief und auf eine Phase der Stabilität oder Besserung einer Dauer von mindestens 30 Tagen folgte. Die Dokumentation eines Schubs setzt voraus, dass der Prüfarzt der Ansicht war, dass mindestens eine neue MS-Läsion vorlag oder eine alte sich vergrößert hatte.</p>
PRISMS	<p>Aggregierte jährliche Schubrate:</p> <p>Gesamtzahl aller bestätigten Schübe aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25</p> <p>Gesamtzahl aller Schübe aller Patienten Mittels Imputation bestimmt anhand der durchschnittlichen Schübe pro Patient</p> <p>Gesamtzahl der Studientage aller Patienten Mittels Imputation bestimmt als (2-Jahres*-Absolvierer) + (1-Jahres*-Verluste für die Nachbeobachtung)</p> <p>Schubdefinition:</p> <p>Ein Schub war gemäß Schumacher et al.⁹ definiert als neues Symptom oder Verschlechterung eines alten Symptoms, das/die auf die MS zurückführbar war, von einer dazu passenden neurologischen Anomalie oder fokalen neurologischen Funktionsstörung begleitet war, über einen Zeitraum von mindestens 24 Stunden anhielt, ohne Fieber verlief und auf eine Phase der Stabilität oder Besserung einer Dauer von mindestens 30 Tagen folgte.</p>
Bornstein et al. (1987)	<p>Aggregierte jährliche Schubrate:</p> <p>Gesamtzahl aller bestätigten Schübe aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25</p> <p>Gesamtzahl aller Schübe aller Patienten Angegeben</p> <p>Gesamtzahl der Studientage aller Patienten In Abbildung angegeben</p> <p>Schubdefinition:</p> <p>Ein Schub war definiert als schnelles Auftreten neuer Symptome oder Verschlechterung vorbestehender Symptome für eine Dauer von 48 Stunden oder länger. Ein Ereignis wurde nur dann als Schub gewertet, wenn die Symptome des Patienten von beobachteten objektiven Veränderungen bei der neurologischen Untersuchung begleitet waren, die einen Anstieg des Punktwerts für eine von acht funktionellen Gruppen der Kurtzke-Skala um mindestens 1 Stufe beinhalteten. Sensorische Symptome, die nicht von objektiven Befunden oder einer</p>

	vorübergehenden neurologischen Verschlechterung begleitet waren, wurden nicht als Schub gewertet.
Aktiv-kontrolliert	
TRANSFORMS	<p>Aggregierte jährliche Schubrate: Gesamtzahl aller bestätigten Schübe (aus CSR) aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25</p> <p>Schubdefinition Ein Schub ist allgemein definiert als neue, verschlechterte oder erneut auftretende neurologische Symptome, die mindestens 30 Tage nach Beginn des vorausgegangenen Schubs auftraten, mindestens 24 Stunden anhielten, ohne Fieber oder Infektion verliefen. Ein bestätigter Schub liegt vor, wenn die allgemeinen Symptome von einem Anstieg um $\geq 0,5$ EDSS-Punkte oder um 1 Punkt in zwei Funktionssystem-Werten oder um 2 Punkte in einem Funktionssystem-Wert (mit Ausnahme von Darm/Blase und psychischen Systemen) begleitet waren.</p>
EVIDENCE	<p>Aggregierte jährliche Schubrate: Gesamtzahl aller bestätigten Schübe aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25</p> <p>Gesamtzahl aller Schübe aller Patienten Mittels Imputation bestimmt anhand der Personenjahre und der berichteten adjustierten jährlichen Schubrate (ARR, nach Zentrum)</p> <p>Gesamtzahl der Studientage aller Patienten Berechnet aus den durchschnittlichen Wochen gemäß Bericht</p> <p>Schubdefinition: Ein Schub war definiert als neues Symptom oder Verschlechterung eines alten Symptoms mit dazu passendem objektivem Befund bei der neurologischen Untersuchung durch den verblindeten Beurteiler, das/die mindestens 24 Stunden anhielt, ohne Fieber verlief und auf einen Zeitraum der klinischen Stabilität oder Besserung von mindestens 30 Tagen folgte. Ein objektiver Befund war definiert als eine Anomalie bei der Untersuchung, die im Einklang mit dem beschriebenen neurologischen Symptom stand. Ein Schub wurde nur dann dokumentiert, wenn der verblindete Beurteiler neue Befunde beschrieb, die im Einklang mit den berichteten Symptomen des Patienten standen und wenn der behandelnde Arzt die Möglichkeit eines Pseudoschubs ausgeschlossen hatte.</p>
INCOMIN	<p>Aggregierte jährliche Schubrate: Gesamtzahl aller bestätigten Schübe aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25</p> <p>Gesamtzahl aller Schübe aller Patienten Mittels Imputation bestimmt anhand der Personenjahre und der jährlichen Schubrate (ARR, Adjustierung unbekannt)</p> <p>Gesamtzahl der Studientage aller Patienten Mittels Imputation bestimmt als (2-Jahres*-Absolvierer) + (1-Jahres*-Verluste für die Nachbeobachtung)</p> <p>Schubdefinition: Schübe waren definiert als neue Symptome oder Verschlechterung alter Symptome mit einer objektiven Änderung des Punktwerts auf der Kurtzke-Skala zu funktionellen Systemen um mindestens 1 Punkt,¹² die mindestens 24 h andauerte(n), ohne Fieber verlief(en) und auf eine Phase der klinischen Stabilität</p>

oder Besserung von mindestens 30 Tagen folgt(en).	
BEYOND	<p>Aggregierte jährliche Schubrate:</p> <p>Gesamtzahl aller bestätigten Schübe aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25</p> <p>Gesamtzahl aller Schübe aller Patienten</p> <p>Mittels Imputation bestimmt anhand der Personenjahre und der jährlichen Schubrate (ARR, Adjustierung unbekannt)</p> <p>Gesamtzahl der Studientage aller Patienten</p> <p>Mittels Imputation bestimmt als (2,75-Jahres*-Absolvierer) + (1-Jahres*-Verluste für die Nachbeobachtung)</p> <p>Schubdefinition:</p> <p>Ein Schub war definiert als neue oder rezidivierende neurologische Störungen, die mindestens 30 Tage nach Beginn des vorherigen Ereignisses auftraten, mindestens 24 Stunden anhielten und ohne Fieber oder Infektion verliefen. Ein neurologisches Ereignis wurde nur dann als Schub gewertet, wenn es mit einem Anstieg eines Punktwertes auf der EDSS-Skala oder den funktionellen Systemskalen einherging – beurteilt durch den verblindeten, beurteilenden Arzt – der für die beschriebenen Symptomen angemessen war.</p>
BECOME	<p>Aggregierte jährliche Schubrate:</p> <p>Gesamtzahl aller bestätigten Schübe aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25</p> <p>Gesamtzahl aller Schübe aller Patienten</p> <p>In der Veröffentlichung angegeben</p> <p>Gesamtzahl der Studientage aller Patienten</p> <p>In der Veröffentlichung angegeben</p> <p>Schubdefinition:</p> <p>Alle neuen oder sich verschlechternden neurologischen Symptome, die 24 Stunden anhielten und nicht durch Fieber oder eine Infektion zu erklären waren, wurden als subjektive Schübe bewertet. Subjektive Schübe, die von einem verblindeten beurteilenden Neurologen auf Grund von sich verschlechternden Punktwerten entweder auf der SNRS-(„Scripps Neurological Rating Scale“)-Skala oder der EDSS-(„Expanded Disability Status Scala“)-Skala bestätigt wurden, wurden als objektive Schübe gewertet. Zur Bestätigung eines Schub musste eine oder mehrere der folgenden Veränderungen gegenüber dem Ausgangszustand vorliegen: 1) Anstieg des EDSS-Gesamtscores um 0,5 Punkte; 2) Anstieg des EDSS-Scores für ein System um 2 Punkte; 3) Anstieg des Punktwerts für zwei oder mehr EDSS-Systeme um 1 Punkt; 4) Abnahme des SNRS-Scores um 7 Punkte.</p>
REGARD	<p>Aggregierte jährliche Schubrate:</p> <p>Gesamtzahl aller bestätigten Schübe aller Patienten in einem Behandlungsarm dividiert durch die Gesamtzahl der Studientage aller Patienten in einem Behandlungsarm und multipliziert mit 365,25</p> <p>Gesamtzahl aller Schübe aller Patienten</p> <p>Mittels Imputation bestimmt auf Grundlage der Gesamtzahl von 395 Schüben und beider jährlicher Schubraten (ARR, adjustiert für das Zentrum)</p> <p>Gesamtzahl der Studientage aller Patienten</p> <p>Mittels Imputation bestimmt auf Grundlage der Gesamtzahl von 395 Schüben und beider ARR (adjustiert für das Zentrum)</p> <p>Schubdefinition:</p>

Ein qualifizierender Schub war definiert als neue oder sich verschlechternde neurologische Symptome ohne Fieber, die über 48 h oder länger anhielten und von einer Änderung des KFS-Scores begleitet waren.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-60: Bewertung des Verzerrungspotenzials für „Jährliche Schubrate“ in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
Placebo-kontrolliert						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
MSCRG	niedrig	ja	ja	nein	nein	niedrig
Johnson et al. (1995)	niedrig	ja	ja	nein	nein	niedrig
IFNB MS Study Group	niedrig	ja	ja	nein	nein	niedrig
PRISMS	niedrig	ja	ja	nein	nein	niedrig
Bornstein et al. (1987)	niedrig	ja	ja	nein	nein	niedrig
Aktiv-kontrolliert						
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
EVIDENCE	niedrig	ja	ja	nein	nein	niedrig
INCOMIN	hoch	nein	ja	nein	nein	hoch
BEYOND	niedrig	ja	ja	nein	nein	niedrig
BECOME	niedrig	ja	ja	nein	nein	niedrig
REGARD	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial für die jährliche Schubrate für alle Studien (außer INCOMIN) als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in allen

Fällen als „niedrig“ eingestuft wurde, jeweils der Endpunkterheber verblindet war, das ITT-Prinzip adäquat umgesetzt wurde und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Faktoren gab.

Für die Studie INCOMIN wurde das Verzerrungspotenzial des Endpunktes als hoch eingestuft, da er nicht verblindet erhoben wurde und das Verzerrungspotenzial auf Studienebene als hoch eingestuft wurde.

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 4-61: Ergebnisse für „Jährliche Schubrate“^a aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1)					
	Fingoli mod	IFN-β-1a i. m.	IFN-β-1a s. c.	IFN-β-1b s. c.	Glatiramer- acetat	Placebo
Placebo-kontrolliert						
FREEDOMS	0,21					0,47
MSCRG		0,67				0,82
Johnson et al. (1995)					0,59	0,84
IFNB MS Study Group				0,84		1,27
PRISMS			0,88			1,32
Bornstein et al. (1987)					0,34	1,38
Aktiv-kontrolliert						
TRANSFORMS	0,21	0,43				
EVIDENCE		0,64	0,54			
INCOMIN		0,70		0,50		
BEYOND				0,36	0,34	
BECOME				0,37	0,33	
REGARD			0,30		0,29	
a: Jährliche Schubrate wie in der Publikation oder im Studienbericht angegeben oder berechnet aus den Publikationsdaten						
Note: die Daten wurden aus Publikationen extrahiert bzw. extrapoliert, eine Angabe von Standardfehler und Konfidenzintervallen ist daher nicht einheitlich möglich. Daher wurde gänzlich auf die Darstellung verzichtet						

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

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

Ergebnisse des indirekten Vergleichs für „Jährliche Schubrate“

Tabelle 4-62: Ergebnisse für „Jährliche Schubrate“ aus RCT für indirekte Vergleiche

	Relative Rate (95 %-KI)
Glatirameracetat vs. Fingolimod	1,411 (1,125; 1,768)
Fingolimod vs. Placebo	0,431 (0,359; 0,519)
Glatirameracetat vs. Placebo	0,608 (0,527; 0,702)

Die Schubrate war unter Glatirameracetat 1,411 mal höher als unter Fingolimod, umgekehrt ausgedrückt war die relative Schubrate unter Fingolimod 0,709. Die Schubrate war damit unter Fingolimod gegenüber Glatirameracetat um etwa 30 % signifikant reduziert. Die durchgeführte Netzwerkanalyse kann für den Wirksamkeitsvergleich zwischen Fingolimod und Glatirameracetat aber nur als Annäherung gelten und ist auch für die gemäß zugelassener Indikation in Frage kommenden Patienten nur mäßig aussagekräftig.

Die relativen Raten aus der Netzwerkanalyse sind für den Vergleich Fingolimod vs. Placebo konsistent mit den Studiendaten (relative Rate 0,447). Für Glatirameracetat vs. Placebo liegen die relativen Raten in den relevanten Studien bei 0,702 (Johnson et al., 1995) und 0,246 (Bornstein et al., 1987) und sind damit zwar nicht in der Größenordnung konsistent, aber in der Richtung. Auffällig ist dabei, dass die Studie von Bornstein et al. besonders deutlich abweicht.

Die Heterogenität wurde in den zugrundeliegenden Modellen mithilfe von Kovariaten untersucht. Die Kovariaten hatten keinen Einfluss auf den Behandlungseffekt. Details sind im Bericht zur Netzwerkanalyse beschrieben (24).

4.3.2.1.3.1.2 Anteil schubfreier Patienten – indirekte Vergleiche aus RCT

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

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

Tabelle 4-63: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studie n	Referenzen Studien	Intervention	IFN- β -1a i. m.	IFN- β -1a s. c.	IFN- β -1b s. c.	Glatiramer acetat	Placebo
Placebo-kontrolliert							
1	FREEDOMS	•					•
1	MSCRG		•				•
2	Johnson et al. (1995) Bornstein et al. (1987)					• •	• •
1	The IFNB MS Study Group				•		•
1	PRISMS			•			•
Aktiv-kontrolliert							
1	TRANSFORMS	•	•				
1	EVIDENCE		•	•			
1	INCOMIN		•		•		
2	BEYOND BECOME				• •	• •	
1	REGARD			•		•	
1	Etemadifar et al. (2006)		•	•	•		

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

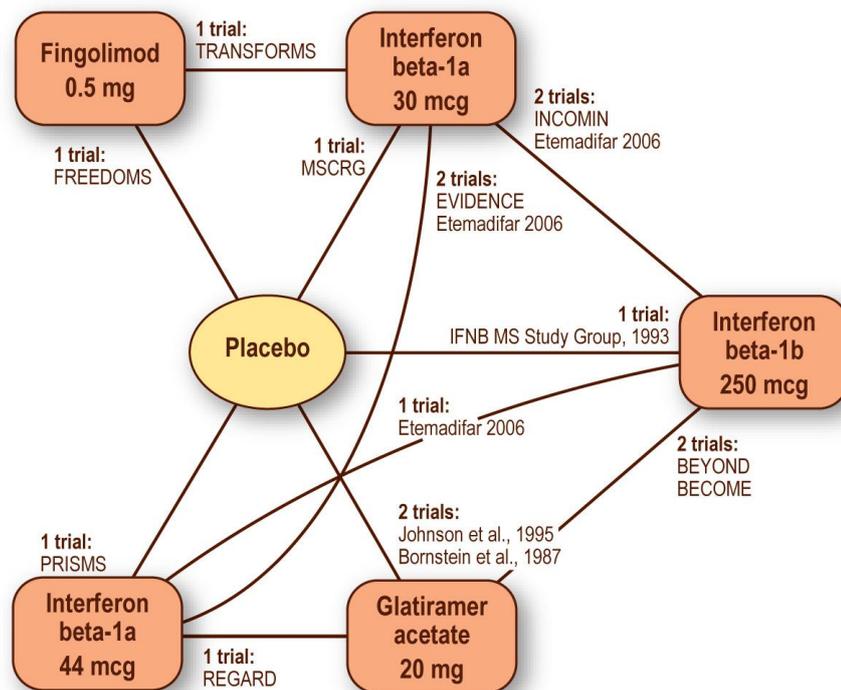


Abbildung 4-27: Netzwerkstruktur für Endpunkt „Anteil schubfreier Patienten“

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

Tabelle 4-64: Operationalisierung von „Anteil schubfreier Patienten“

Studie	Operationalisierung
Placebo-kontrolliert	
FREEDOMS	<p>Anteil schubfreier Patienten zu Monat 24 (Kaplan-Meier Schätzer aus „Zeit bis zum ersten bestätigten Schub“)</p> <p>Schubdefinition:</p> <p>Ein Schub ist allgemein definiert als neue, verschlechterte oder erneut auftretende neurologische Symptome, die mindestens 30 Tage nach Beginn des vorausgegangenen Schubs auftraten, mindestens 24 Stunden anhielten, ohne Fieber oder Infektion verliefen.</p> <p>Ein bestätigter Schub liegt vor, wenn die allgemeinen Symptome von einem Anstieg um $\geq 0,5$ EDSS-Punkte oder um 1 Punkt in zwei Funktionssystem-Werten oder um 2 Punkte in einem Funktionssystem-Wert (mit Ausnahme von Darm/Blase und psychischen Systemen) begleitet waren.</p>
MSCRG	<p>Anteil schubfreier Patienten zu Monat 24</p> <p>Schubdefinition:</p> <p>Schübe im Studienverlauf waren definiert als neue neurologische Symptome oder Verschlechterung bestehender neurologischer Symptome, die über mindestens 48 Stunden anhielt(en), bei einem Patienten, der in den vorangegangenen 30 Tagen klinisch stabil war oder eine Besserung gezeigt hatte, begleitet von einer objektiven Änderung bei der neurologischen Untersuchung (Verschlechterung auf der EDSS-Skala um 0,5 Punkte oder Verschlechterung des Punktwerts für das pyramidale, zerebellare, den Hirnstamm betreffende oder visuelle funktionelle System um mindestens 1,0 Punkte)</p>
Johnson et al. (1995)	<p>Anteil schubfreier Patienten zu Monat 24</p> <p>Schubdefinition:</p> <p>Ein Schub war definiert als Auftreten oder erneutes Auftreten einer oder mehrerer neurologischer Störungen, die über mindestens 48 Stunden anhielten und denen unmittelbar eine Phase der relativen Stabilität oder Verbesserung des neurologischen Zustands einer Dauer von mindestens 30 Tagen vorausgegangen war. Zur Bestätigung eines Schubs wurde eine objektive Veränderung bei der neurologischen Untersuchung verwendet, die einem Anstieg um mindestens eine halbe Stufe im EDSS-Score, 2 Punkte bei einem der sieben funktionellen Systeme oder 1 Punkt bei zwei oder mehr funktionellen Systemen entsprach.</p>
IFNB MS Study Group	<p>Anteil schubfreier Patienten zu Monat 24</p> <p>Schubdefinition:</p> <p>Ein Schub war definiert als neues Symptom oder Verschlechterung eines alten Symptoms, das/die auf die MS zurückzuführen war, von einer dazu passenden neuen neurologischen Störung begleitet war, über einen Zeitraum von mindestens 24 Stunden anhielt, ohne Fieber verlief und auf eine Phase der Stabilität oder Besserung von einer Dauer von mindestens 30 Tagen folgte. Die Dokumentation eines Schubs setzte voraus, dass der Prüfarzt der Ansicht war, dass mindestens eine neue MS-Läsion vorlag oder eine alte sich vergrößert hatte.</p>
PRISMS	<p>Anteil schubfreier Patienten zu Monat 24</p> <p>Schubdefinition:</p> <p>Ein Schub war gemäß Schumacher et al.⁹ definiert als neues Symptom oder Verschlechterung eines alten Symptoms, das/die auf die MS zurückführbar war, von einer dazu passenden neurologischen Anomalie oder fokalen neurologischen Funktionsstörung begleitet war, über einen Zeitraum von</p>

	mindestens 24 Stunden anhielt, ohne Fieber verlief und auf eine Phase der Stabilität oder Besserung einer Dauer von mindestens 30 Tagen folgte.
Bornstein et al. (1987)	<p>Anteil schubfreier Patienten zu Monat 24</p> <p>Schubdefinition:</p> <p>Ein Schub war definiert als schnelles Auftreten neuer Symptome oder Verschlechterung vorbestehender Symptome mit einer Dauer von 48 Stunden oder länger. Ein Ereignis wurde nur dann als Schub gewertet, wenn die Symptome des Patienten von beobachteten objektiven Veränderungen bei der neurologischen Untersuchung begleitet waren, die einen Anstieg des Punktwerts bei einer von acht funktionellen Gruppen der Kurtzke-Skala um mindestens 1 Stufe beinhalteten. Sensorische Symptome, die nicht von objektiven Befunden oder einer vorübergehenden neurologischen Verschlechterung begleitet waren, wurden nicht als Schub gewertet.</p>
Aktiv-kontrolliert	
TRANSFORMS	<p>Anteil schubfreier Patienten zu Monat 12 (Kaplan-Meier Schätzer aus „Zeit bis zum ersten bestätigten Schub“)</p> <p>Schubdefinition:</p> <p>Ein Schub ist allgemein definiert als neue, verschlechterte oder erneut auftretende neurologische Symptome, die mindestens 30 Tage nach Beginn des vorausgegangenen Schubs auftraten, mindestens 24 Stunden anhielten, ohne Fieber oder Infektion verliefen.</p> <p>Ein bestätigter Schub liegt vor, wenn die allgemeinen Symptome von einem Anstieg um $\geq 0,5$ EDSS-Punkte oder um 1 Punkt in zwei Funktionssystem-Werten oder um 2 Punkte in einem Funktionssystem-Wert (mit Ausnahme von Darm/Blase und psychischen Systemen) begleitet waren.</p>
EVIDENCE	<p>Anteil schubfreier Patienten zu Monat 12</p> <p>Schubdefinition:</p> <p>Ein Schub war definiert als neues Symptom oder Verschlechterung eines alten Symptoms mit dazu passendem objektivem Befund bei der neurologischen Untersuchung durch den verblindeten Beurteiler, das/die mindestens 24 Stunden anhielt, ohne Fieber verlief und auf einen Zeitraum der klinischen Stabilität oder Besserung von mindestens 30 Tagen folgte. Ein objektiver Befund war definiert als eine Anomalie bei der Untersuchung, die im Einklang mit dem beschriebenen neurologischen Symptom stand. Ein Schub wurde nur dann dokumentiert, wenn der verblindete Beurteiler neue Befunde beschrieb, die im Einklang mit den berichteten Symptomen des Patienten standen und wenn der behandelnde Arzt die Möglichkeit eines Pseudoschubs ausgeschlossen hatte.</p>
INCOMIN	<p>Anteil schubfreier Patienten zu Monat 24</p> <p>Schubdefinition:</p> <p>Schübe waren definiert als neue Symptome oder Verschlechterung alter Symptome mit einer objektiven Änderung des Punktwerts auf der Kurtzke-Skala zu funktionellen Systemen um mindestens 1 Punkt,¹² die mindestens 24 h andauerte(n), ohne Fieber verlief(en) und auf eine Phase der klinischen Stabilität oder Besserung von mindestens 30 Tagen folgt(en).</p>
BEYOND	<p>Anteil schubfreier Patienten zu Monat 24</p> <p>Schubdefinition:</p> <p>Ein Schub war definiert als neue oder rezidivierende neurologische Störungen, die mindestens 30 Tage nach Beginn des vorherigen Ereignisses auftraten, mindestens 24 Stunden anhielten und ohne Fieber oder Infektion verliefen. Ein neurologisches Ereignis wurde nur dann als Schub gewertet, wenn es mit einem Anstieg eines Punktwertes auf der EDSS-Skala oder den funktionellen Systemskalen einherging – beurteilt durch den verblindeten, beurteilenden Arzt</p>

– der für die beschriebenen Symptome angemessen war.	
BECOME	<p>Anteil schubfreier Patienten zu Monat 24</p> <p>Schubdefinition:</p> <p>Alle neuen oder sich verschlechternden neurologischen Symptome, die 24 Stunden anhielten und nicht durch Fieber oder eine Infektion zu erklären waren, wurden als subjektive Schübe bewertet. Subjektive Schübe, die von einem verblindeten beurteilenden Neurologen auf Grund von sich verschlechternden Punktwerten entweder auf der SNRS-(„Scripps Neurological Rating Scale“)-Skala oder der EDSS-(„Expanded Disability Status Scala“)-Skala bestätigt wurden, wurden als objektive Schübe gewertet. Zur Bestätigung eines Schub musste eine oder mehrere der folgenden Veränderungen gegenüber dem Ausgangszustand vorliegen: 1) Anstieg des EDSS-Gesamtscores um 0,5 Punkte; 2) Anstieg des EDSS-Scores für ein System um 2 Punkte; 3) Anstieg des Punktwerts für zwei oder mehr EDSS-Systeme um 1 Punkt; 4) Abnahme des SNRS-Scores um 7 Punkte.</p>
REGARD	<p>Anteil schubfreier Patienten zu Monat 24</p> <p>Schubdefinition:</p> <p>Ein qualifizierender Schub war definiert als neue oder sich verschlechternde neurologische Symptome ohne Fieber, die über 48 h oder länger anhielten und von einer Änderung des KFS-Scores begleitet waren.</p>
Etemadifar et al. (2006)	<p>Anteil schubfreier Patienten zu Monat 24</p> <p>Schubdefinition:</p> <p>Ein akuter Schub war definiert als Auftreten eines neuen neurologischen Symptoms oder starke Verschlechterung eines vorbestehenden Symptoms, das/die über mindestens 24 Stunden anhielt und zu einer Verschlechterung des EDSS-Scores um 1 Punkt führte.</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-65: Bewertung des Verzerrungspotenzials für „Anteil schubfreier Patienten“ in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
Placebo-kontrolliert						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
MSCRG	niedrig	ja	ja	nein	nein	niedrig
Johnson et al. (1995)	niedrig	ja	ja	nein	nein	niedrig
IFNB MS Study Group	niedrig	ja	ja	nein	nein	niedrig
PRISMS	niedrig	ja	ja	nein	nein	niedrig
Bornstein et al. (1987)	niedrig	ja	ja	nein	nein	niedrig
Aktiv-kontrolliert						
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
EVIDENCE	niedrig	ja	ja	nein	nein	niedrig
INCOMIN	hoch	nein	ja	nein	nein	hoch
BEYOND	niedrig	ja	ja	nein	nein	niedrig
BECOME	niedrig	ja	ja	nein	nein	niedrig
REGARD	niedrig	ja	ja	nein	nein	niedrig
Etemadifar et al. (2006)	niedrig	ja	nein	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial für Anteil schubfreier Patienten für alle Studien (außer INCOMIN) als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in allen Fällen als „niedrig“ eingestuft wurde, jeweils der Endpunkterheber verblindet war, das ITT-Prinzip adäquat umgesetzt wurde (außer Etemadifar) und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Faktoren gab. Das Verzerrungspotenzial für die Studie Etemadifar et al. (2006) wurde trotzdem auf Ebene dieses Endpunktes als niedrig eingestuft, da die Studie als Sicherheitsstudie ausgerichtet war und daher keine ITT-Population definiert wurde.

Für die Studie INCOMIN wurde das Verzerrungspotenzial des Endpunktes als hoch eingestuft, da er nicht verblindet erhoben wurde und das Verzerrungspotential auf Studienebene als hoch eingestuft wurde.

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 4-66: Ergebnisse für „Anteil schubfreier Patienten“^{ac} aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1)					
	Fingolimod	IFN-β-1a i. m.	IFN-β-1a s. c.	IFN-β-1b s. c.	Glatiramer acetat	Placebo
Placebo-kontrolliert						
FREEDOMS	0,70					0,46
MSCRG		0,38				0,26
Johnson et al. (1995)					0,34	0,27
IFNB MS Study Group				0,31		0,16
PRISMS			0,32			0,16
Bornstein et al. (1987)					0,56	0,26
Aktiv-kontrolliert						
TRANSFORMS	0,83	0,70				
EVIDENCE		0,52	0,62			
INCOMIN		0,36		0,51		
BEYOND				0,59	0,59	
BECOME				0,53	0,72	
REGARD			0,67		0,65	
Etemadifar et al. (2006)		0,20	0,57	0,43		
a: wie in der Publikation angegeben oder daraus ermittelt Anmerkung: die Daten wurden aus Publikationen extrahiert bzw. extrapoliert, eine Angabe von Standardfehler und Konfidenzintervallen ist daher nicht einheitlich möglich. Daher wurde gänzlich auf die Darstellung verzichtet						

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

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

Ergebnisse des indirekten Vergleichs für „Anteil schubfreier Patienten“

Tabelle 4-67: Ergebnisse für „Anteil schubfreier Patienten“ aus RCT für indirekte Vergleiche

	Relative Rate (95 %-KI)
Glatirameracetat vs. Fingolimod	1,034 (0,880; 1,214)
Fingolimod vs. Placebo	1,584 (1,398; 1,795)
Glatirameracetat vs. Placebo	1,638 (1,375; 1,950)

Der Anteil schubfreier Patienten ist unter Fingolimod und Glatirameracetat vergleichbar. Zusammen mit der Aussage zur reduzierten Schubrate unter Fingolimod bedeutet das, dass zwar gleichviele Patienten Schübe haben, aber die Zahl der Schübe pro Patient unter Fingolimod reduziert wurde. Es gilt aber auch hier, dass die Interpretation der Daten mit Vorsicht erfolgen muss, da die Daten auf einem nur bedingt für die relevanten Patientenpopulationen aussagekräftigen Vergleich beruhen.

Die relativen Raten aus der Netzwerkanalyse sind für den Vergleich Fingolimod vs. Placebo konsistent mit den Studiendaten (RR 1,521). Für Glatirameracetat vs. Placebo liegen die relativen Raten in den relevanten Studien bei 1,259 (Johnson et al., 1995) und 2,154 (Bornstein et al., 1987) und sind damit zwar nicht in der Größenordnung konsistent, aber in der Richtung. Auffällig ist dabei, dass die Studie von Bornstein et al. besonders deutlich abweicht.

Die Heterogenität wurde in den zugrundeliegenden Modellen mithilfe von Kovariaten untersucht. Die Kovariaten hatten keinen Einfluss auf den Behandlungseffekt. Details sind im Bericht zur Netzwerkanalyse beschrieben (24).

4.3.2.1.3.1.3 Nach 3 Monaten bestätigte Behinderungsprogression – indirekte Vergleiche aus RCT

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

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

Tabelle 4-68: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Referenzen Studien	Intervention	IFN- β -1a i. m.	IFN- β -1a s. c.	IFN- β -1b s. c.	Glatirameracetat	Placebo
Placebo-kontrolliert							
1	FREEDOMS	•					•
2	Johnson et al. (1995) Bornstein et al. (1987)					• •	• •
1	The IFNB MS Study Group				•		•
Aktiv-kontrolliert							
1	BEYOND				•	•	

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

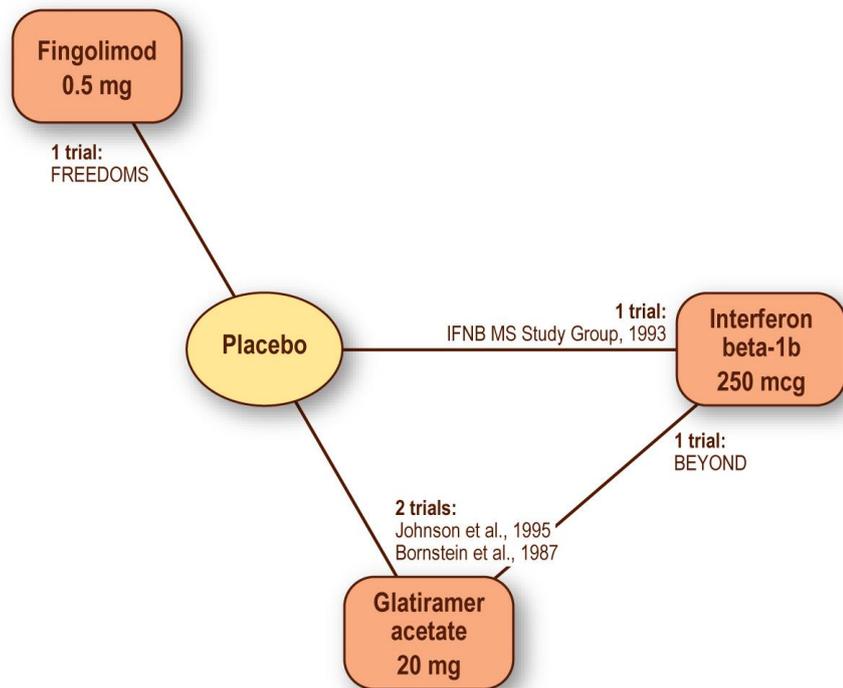


Abbildung 4-28: Netzwerkstruktur für Endpunkt „Behinderungsprogression“

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

Tabelle 4-69: Operationalisierung von „nach 3 Monaten bestätigte Behinderungsprogression“

Studie	Operationalisierung
Placebo-kontrolliert	
FREEDOMS	Risiko einer nach 3 Monaten bestätigten Behinderungsprogression zu Monat 24 ^a (Kaplan-Meier-Schätzer) zu Monat 24 Die Behinderungsprogression war definiert als eine Erhöhung des EDSS-Werts um 1 Punkt (um 0,5 Punkte bei einem Baseline-EDSS-Wert $\geq 5,5$), die 3 Monate später ohne Vorliegen eines Schubs bestätigt wurde.
Johnson et al. (1995)	Risiko einer nach 3 Monaten bestätigten Behinderungsprogression zu Monat 24, berechnet aus den Patientenzahlen (mittels Imputation aus Prozentzahlen) Eine anhaltende Krankheitsprogression war definiert als Zunahme des EDSS-Werts um mindestens einen Punkt, die mindestens drei Monate anhielt.
Bornstein et al. (1987)	Risiko einer nach 3 Monaten bestätigten Behinderungsprogression zu Monat 24, berechnet aus den angegebenen Patientenzahlen Eine Progression war definiert als Zunahme des Krutzke-Scores um mindestens eine Einheit, die mindestens 3 Monate anhielt.
The IFNB MS Study Group	Risiko einer nach 3 Monaten bestätigten Behinderungsprogression zu Monat 36, berechnet aus den angegebenen Patientenzahlen Eine bestätigte Behinderung war definiert als 2 aufeinanderfolgende identische Bestimmungen des EDSS-Werts im Abstand von 90 Tagen, die einen Anstieg des Punktwerts um 1,0 Punkt jeweils gegenüber dem Ausgangswert zeigten.
Aktiv-kontrolliert	
BEYOND	Risiko einer nach 3 Monaten bestätigten Behinderungsprogression zu Monat 24, berechnet aus den angegebenen Patientenzahlen Eine Progression auf der EDSS-Skala war definiert als Änderung des Punktwerts um mindestens 1 Punkt, die mindestens 3 Monate anhielt.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-70: Bewertung des Verzerrungspotenzials für „nach 3 Monaten bestätigte Behinderungsprogression“ in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
Placebo-kontrolliert						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
Johnson et al. (1995)	niedrig	ja	ja	nein	nein	niedrig
Bornstein et al. (1987)	niedrig	ja	ja	nein	nein	niedrig
The IFNB MS Study Group	niedrig	ja	ja	nein	nein	niedrig
Aktiv-kontrolliert						
BEYOND	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial für Anteil behinderungsprogressionsfreier Patienten für alle Studien als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in allen Fällen als „niedrig“ eingestuft wurde, jeweils der Endpunkterheber verblindet war, das ITT-Prinzip adäquat umgesetzt wurde und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Faktoren gab.

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 4-71: Ergebnisse für “Risiko einer nach 3 Monaten bestätigten Behinderungsprogression^{ac}“ aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1)					
	Fingolimod	IFN-β-1a i. m.	IFN-β-1a s. c.	IFN-β-1b s. c.	Glatiramer acetat	Placebo
Placebo-kontrolliert						
FREEDOMS	0,18					0,24
Johnson et al. 1995					0,22	0,25
Bornstein et al. 1987					0,09	0,48
The IFNB MS Study Group				0,20		0,28
Aktiv-kontrolliert						
BEYOND					0,21	0,22
a: wie in der Publikation oder im Studienbericht angegeben oder berechnet aus den Publikationsdaten Anmerkung: die Daten wurden aus Publikationen extrahiert bzw. extrapoliert, eine Angabe von Standardfehler und Konfidenzintervallen ist daher nicht einheitlich möglich. Daher wurde gänzlich auf die Darstellung verzichtet						

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

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

Ergebnisse des indirekten Vergleichs für „Nach 3 Monaten bestätigte Behinderungsprogression“

Tabelle 4-72: Ergebnisse für „Nach 3 Monaten bestätigte Behinderungsprogression“ aus RCT für indirekte Vergleiche

	Relative Rate (95 %-KI)
Glatirameracetat vs. Fingolimod	1,175 (0,666; 2,072)
Fingolimod vs. Placebo	0,688 (0,405; 1,169)
Glatirameracetat vs. Placebo	0,808 (0,519; 1,259)

Die Rate der nach 3 Monaten bestätigten Behinderungsprogression ist unter Glatirameracetat und Fingolimod vergleichbar. Die durchgeführte Netzwerkanalyse kann nur als Annäherung gelten und ist für die gemäß zugelassener Indikation in Frage kommenden Patienten nur mäßig aussagekräftig.

Die relativen Raten aus der Netzwerkanalyse sind für den Vergleich Fingolimod vs. Placebo konsistent mit den Studiendaten (RR 0,75). Für Glatirameracetat vs. Placebo liegen die relativen Raten in den relevanten Studien bei 0,88 (Johnson et al., 1995) und 0,417 (Bornstein et al., 1987) und sind damit zwar nicht in der Größenordnung konsistent, aber in der Richtung. Auffällig ist dabei, dass die Studie von Bornstein et al. besonders deutlich abweicht.

Die Heterogenität wurde in den zugrundeliegenden Modellen mithilfe von Kovariaten untersucht. Die Kovariaten hatten keinen Einfluss auf den Behandlungseffekt. Details sind im Bericht zur Netzwerkanalyse beschrieben (24).

4.3.2.1.3.1.4 Behandlungsabbrüche aufgrund von UE – indirekte Vergleiche aus RCT

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

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

Tabelle 4-73: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Referenzen Studien	Intervention	IFN-β-1a i. m.	IFN-β-1a s. c.	IFN-β-1b s. c.	Glatiramer acetat	Placebo
Placebo-kontrolliert							
1	FREEDOMS	•					•
1	The IFNB MS Study Group				•		•
1	PRISMS			•			•
Aktiv-kontrolliert							
1	TRANSFORMS	•	•				
1	EVIDENCE		•	•			
1	INCOMIN		•		•		
1	BEYOND				•	•	
1	REGARD			•		•	

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

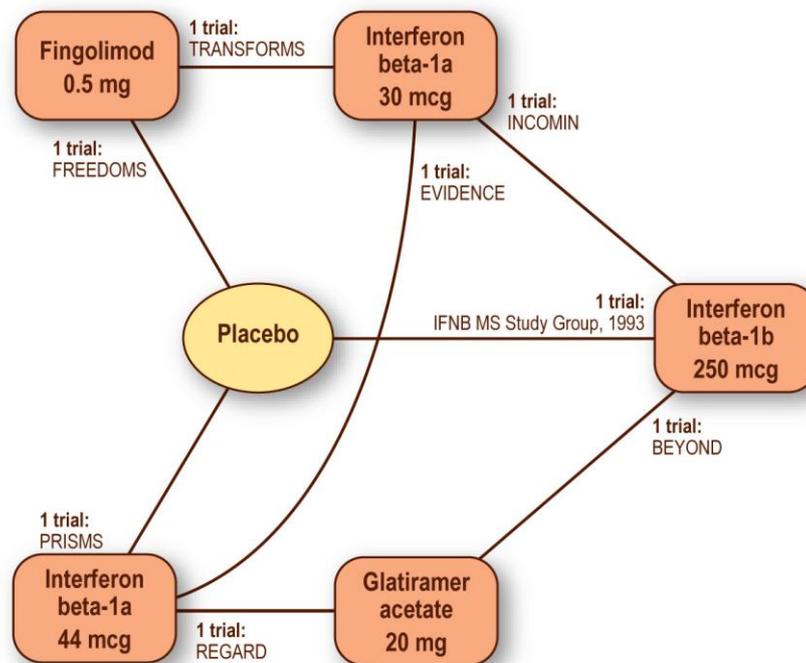


Abbildung 4-29: Netzwerkstruktur für Endpunkt „Behandlungsabbrüche aufgrund von UE“

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 von „Behandlungsabbrüche aufgrund von UE“

Studie	Operationalisierung
Placebo-kontrolliert	
FREEDOMS	Anzahl an Patienten mit unerwünschten Ereignissen in 24 Monaten, die zum Absetzen der Studienmedikation geführt haben
The IFNB MS Study Group	Anzahl an Patienten mit unerwünschten Ereignissen in 24 Monaten, die zum Absetzen der Studienmedikation geführt haben
PRISMS	Anzahl an Patienten mit unerwünschten Ereignissen in 24 Monaten, die zum Absetzen der Studienmedikation geführt haben
Aktiv-kontrolliert	
TRANSFORMS	Anzahl an Patienten mit unerwünschten Ereignissen in 12 Monaten, die zum Absetzen der Studienmedikation geführt haben
EVIDENCE	Anzahl an Patienten mit unerwünschten Ereignissen in 12 Monaten, die zum Absetzen der Studienmedikation geführt haben
INCOMIN	Anzahl an Patienten mit unerwünschten Ereignissen in 24 Monaten, die zum Absetzen der Studienmedikation geführt haben
BEYOND	Anzahl an Patienten mit unerwünschten Ereignissen in 24 Monaten, die zum Absetzen der Studienmedikation geführt haben
REGARD	Anzahl an Patienten mit unerwünschten Ereignissen in 24 Monaten, die zum Absetzen der Studienmedikation geführt haben

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-75: Bewertung des Verzerrungspotenzials für „Behandlungsabbrüche aufgrund von UE“ in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
Placebo-kontrolliert						
FREEDOMS	niedrig	ja	ja	nein	nein	niedrig
The IFNB MS Study Group	niedrig	ja	ja	nein	nein	niedrig
PRISMS	niedrig	ja	ja	nein	nein	niedrig
Aktiv-kontrolliert						
TRANSFORMS	niedrig	ja	ja	nein	nein	niedrig
EVIDENCE	niedrig	ja	ja	nein	nein	niedrig
INCOMIN	hoch	nein	ja	nein	nein	hoch
BEYOND	niedrig	nein	ja	nein	nein	niedrig
REGARD	niedrig	unklar	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Insgesamt wird das Verzerrungspotenzial für Anteil schubfreier Patienten für alle Studien (außer INCOMIN) als „niedrig“ bewertet, da das Verzerrungspotenzial auf Studienebene in allen Fällen als „niedrig“ eingestuft wurde, jeweils der Endpunkterheber verblindet war (außer BEYOND), das ITT-Prinzip adäquat umgesetzt wurde und es weder Hinweise auf ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Faktoren gab. Das Verzerrungspotenzial für die Studie BEYOND wurde auf Ebene dieses Endpunktes trotzdem als niedrig bewertet, da die Verblindung nicht relevant für die Erfassung der Behandlungsabbrüche aufgrund von UE ist.

Für die Studie INCOMIN wurde das Verzerrungspotenzial des Endpunktes als hoch eingestuft, da das Verzerrungspotenzial auf Studienebene als hoch eingestuft wurde.

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 4-76: Ergebnisse für „Behandlungsabbrüche aufgrund von UE^{acc}“ aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1)					
	Fingolimod n (%)	IFN-β-1a i. m. n (%)	IFN-β-1a s. c. n (%)	IFN-β-1b s. c. n (%)	Glatiramer- acetat n (%)	Placebo n (%)
Placebo-kontrolliert						
FREEDOMS	15 (0,035)					24 (0,057)
The IFNB MS Study Group				10 (0,018)		1 (0,008)
PRISMS			9 (0,049)			2 (0,011)
Aktiv-kontrolliert						
TRANSFORMS	16 (0,037)	12 (0,028)				
EVIDENCE		14 (0,041)	16 (0,047)			
INCOMIN		1 (0,011)		5 (0,052)		
BEYOND				13 (0,015)	8 (0,018)	
REGARD			23 (0,060)		19 (0,051)	
a: wie in der Publikation angegeben oder daraus ermittelt; Anzahl Patienten mit Behandlungsabbruch aufgrund von UE (in Klammern daraus ermitteltes Risiko eines UE-bedingten Behandlungsabbruchs)						
Note: die Daten wurden aus Publikationen extrahiert bzw. extrapoliert, eine Angabe von Standardfehler und Konfidenzintervallen ist daher nicht einheitlich möglich. Daher wurde gänzlich auf die Darstellung verzichtet						

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

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

Ergebnisse des indirekten Vergleichs für „Unerwünschte Ereignisse“

Tabelle 4-77: Ergebnisse für „Behandlungsabbrüche aufgrund unerwünschte Ereignisse“ aus RCT für indirekte Vergleiche

	Relative Rate (95 %-KI)
Glatirameracetat vs. Fingolimod	1,723 (0,494; 6,009)
Fingolimod vs. Placebo	0,924 (0,376; 2,270)
Glatirameracetat vs. Placebo	1,592 (0,465; 5,456)

Die Rate an Behandlungsabbrüchen aufgrund von UE ist unter Glatirameracetat um das 1,7 fache erhöht. Fingolimod bietet tendenziell einen Sicherheitsvorteil gegenüber Glatirameracetat.

Die relativen Raten aus der Netzwerkanalyse sind für den Vergleich Fingolimod vs. Placebo mit den Studiendaten (RR 0,614) in der Richtung konsistent. Für Glatirameracetat vs. Placebo kann keine Beurteilung erfolgen, da keine Daten diesbezüglich aus den Publikationen extrahiert werden konnten.

Die Heterogenität wurde in den zugrundeliegenden Modellen mithilfe von Kovariaten untersucht. Die Kovariaten hatten keinen Einfluss auf den Behandlungseffekt. Details sind im Bericht zur Netzwerkanalyse beschrieben (24).

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

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.

Subgruppenanalysen wurden im indirekten Vergleich nicht durchgeführt, da die dafür herangezogenen Publikationen keine ausreichende Datengrundlage bieten.

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. Benennen Sie

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

Nicht zutreffend.

4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien

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

Beschreiben Sie die Verzerrungsaspekte der nicht randomisierten vergleichenden Studie auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-78: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Andere Aspekte, die Verzerrungen verursachen können
			Patient	Behandler		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Nicht zutreffend.

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

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

Tabelle 4-79: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Beschreiben Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. 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-G.

Tabelle 4-80: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Ver- zerrungspotenzial be- einflussende Punkte
<Studie 1>	ja / nein / unklar	ja / nein / unklar	ja / nein / unklar	ja / nein

1) Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar.

Nicht zutreffend.

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

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.

Nicht zutreffend.

4.3.2.3 Weitere Untersuchungen

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

4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

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

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

Nicht zutreffend.

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

Charakterisieren Sie nachfolgend die weiteren Untersuchungen und bewerten Sie deren Verzerrungspotenzial.

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

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

Nicht zutreffend.

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen

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

Tabelle 4-81: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

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

Nicht zutreffend.

Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung.

Nicht zutreffend.

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

4.3.2.3.2 Subgruppenanalysen – Weitere Untersuchungen

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

Nicht zutreffend.

4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen

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

Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.

Im indirekten Vergleich zeigte sich Fingolimod gegenüber Glatirameracetat hinsichtlich der Schubrate signifikant und hinsichtlich der Behandlungsabbrüche aufgrund von UE tendenziell überlegen:

- Die Schubrate war unter Fingolimod gegenüber Glatirameracetat um 30 % signifikant reduziert.
- Der Anteil schubfreier Patienten ist unter Fingolimod und Glatirameracetat vergleichbar.
- Die Rate der nach 3 Monaten bestätigten Behinderungsprogression ist unter Glatirameracetat und Fingolimod vergleichbar.
- Die Rate an Behandlungsabbrüchen aufgrund von UE ist unter Glatirameracetat um das 1,7-fache erhöht. Fingolimod bietet tendenziell einen Sicherheitsvorteil gegenüber Glatirameracetat.

Eine annähernd gleiche Anzahl an Patienten ohne Schübe bei deutlich reduzierter Schubrate unter Fingolimod bedeutet, dass zwar gleichviele Patienten Schübe haben, aber die Zahl der Schübe pro Patient unter Fingolimod reduziert wurde.

Die Aussagen aus dem indirekten Vergleich müssen mit Vorsicht interpretiert werden. Das Netzwerk enthält Studien, die aufgrund der Zeit der Durchführung mit unterschiedlichen Diagnosekriterien gearbeitet haben. Ein Teil der Studien (vor 2001) arbeitete noch mit den Poser-Kriterien. Spätere Studien legen die McDonald-Kriterien zugrunde (s. auch Modul 3, Abschnitt 3.2 zu den Kriterien). Die resultierenden Studienpopulationen können sich daher hinsichtlich der Krankheitsschwere unterscheiden. Der Einfluss auf das Ergebnis der jeweiligen Studie beziehungsweise des indirekten Vergleichs lässt sich schwer abschätzen. Laut Klawiter et al scheint es aber eine Auswirkung auf die Größenordnungen des Effektmaßes zu geben (84).

Ein weiterer Unterschied zwischen den Studien besteht in der Vortherapie. Einige Studien schließen Vortherapie aus, andere nicht (s. Abschnitt 4.2.5.6). Auch die jeweils erlaubten Vortherapien beziehungsweise die Zeitspanne, die vor Studienbeginn therapiefrei sein musste unterscheidet sich zwischen den Studien. Die Studienpopulationen sind daher sehr unterschiedlich, was möglicherweise auch einen verzerrenden Effekt haben kann. Insgesamt ist der hier erbrachte indirekte Vergleich auf Basis einer Netzwerkanalyse jedoch die bestverfügbare Evidenz für die Frage nach dem Zusatznutzen von Fingolimod gegenüber Glatirameracetat. Ein Bezug zur relevanten Patientengruppe (IFN Non-Responder) war mit der verfügbaren Datengrundlage nicht möglich.

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 medizinischen Nutzens und Zusatznutzens im direkten und indirekten Vergleich wurden ausschließlich randomisierte RCT (Evidenzstufe 1b) eingeschlossen. Die Studien sind für die zu bewertende Indikation und die relevanten Patientengruppen aufgrund der Population (RRMS-Erkrankung, Krankheitsaktivität und Altersstruktur), Fallzahl, Studiendauer, Art der Intervention und der Wahl der Endpunkte relevant und erlauben eine valide Beurteilung des Nutzens und Zusatznutzens. Auch im Hinblick auf das Verzerrungspotenzial zeigt sich eine sehr hohe Studienqualität und Validität der herangezogenen Studien und deren Endpunkte.

Die Studien enthielten patientenrelevante Endpunkte zur Morbidität (Schübe, Behinderungsprogression, UE) und gesundheitsbezogenen Lebensqualität. Studien mit Endpunkten zur Mortalität waren nicht verfügbar. Die Studiendauer der ausgewählten RCT beträgt mindestens 12 Monate. Die Dauer begründet sich auf den schubbezogenen und behinderungsprogressions-bezogenen Endpunkten, die erst ab ausreichender Studiendauer aussagekräftig sind. Um den medizinischen Nutzen von Fingolimod zu zeigen, wurde gegen Placebo getestet; der medizinische Zusatznutzen zeigte sich im Vergleich zur zweckmäßigen Vergleichstherapie IFN- β -1a i. m. beziehungsweise Glatirameracetat.

Die Aussagekraft der oben genannten RCT ist somit als sehr hoch einzuschätzen und damit für eine valide Beurteilung des medizinischen Nutzens und Zusatznutzens uneingeschränkt geeignet. Die verwendeten Endpunkte sind außerdem, wie in Abschnitt 4.2.5.1 ausführlich erläutert, patientenrelevant.

Der indirekte Vergleich bietet für die Fragestellung des Zusatznutzens gegenüber Glatirameracetat die best-verfügbare Evidenz, da keine direkten Vergleichsstudien existieren. Die Methodik des indirekten Vergleichs folgt den Empfehlungen des G-BA und IQWiG und damit dem aktuellen Stand der Wissenschaft und wird daher als valide angesehen.

Die Ergebnisse der einzelnen Endpunkte der direkten Vergleiche werden für die Gesamtpopulation sowie für die relevanten Patientenpopulationen, wie in der Fragestellung formuliert, gegenübergestellt. Eine Meta-Analyse der Daten war nicht möglich, da nur eine Placebo-kontrollierte und eine aktiv-kontrollierte Studie vorliegen. Für den indirekten Vergleich wurden Endpunkte gewählt, die für alle relevanten Studien aus den jeweiligen Publikationen beziehungsweise Berichten auswertbar waren. Die Ergebnisse der relevanten Studien wurden mittels Meta-Analysen nach Wirkstoff zusammengefasst und dann im indirekten Vergleich mit Fingolimod dargestellt.

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. 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 Fingolimod ergibt sich in der Zusammenschau der Ergebnisse ein medizinischer Nutzen und erheblicher Zusatznutzen gegenüber den zweckmäßigen Vergleichstherapien. Die relevanten Patientenpopulationen wurden soweit möglich abgebildet, eine Unterscheidung der IFN Non-Responder in therapietreue und nicht-therapietreue Patienten war mit der verfügbaren Datenlage nicht möglich (s. Abschnitt 4.2.5.1). Die ausgewerteten Patientenpopulationen enthalten jeweils beide Gruppen und sind daher für beide aussagekräftig.

Tabelle 4-82: Bewertung des Zusatznutzens innerhalb der relevanten Patientenpopulationen

	Gesamtpopulation gegen IFN-β-1a i. m.	Anwendungsgebiet a2 gegen IFN-β-1a i. m. (dargestellt über Population IFN Non-Responder Def I und Def II)	Anwendungsgebiet b gegen IFN-β-1a i. m. (Schnell fortschreitende schwere RRMS)	Anwendungsgebiet a1 gegen Glatirameracetat (über indirekten Vergleich) ^b

Reduziert die jährliche Schubrate	✓✓	✓✓	✓	✓✓
Reduziert das Schubrisiko bzw erhöht den Anteil schubfreier Patienten	✓✓	✓✓	✓	=
Reduziert die Anzahl der schweren Schübe	✓ ^a	✓ ^a	✓ ^a	Nicht bestimmbar
reduziert die Anzahl der hospitalisierungspflichtigen Schübe	✓	✓	✓	Nicht bestimmbar
Reduziert die Anzahl steroidpflichtiger Schübe	✓	✓	✓	Nicht bestimmbar
Reduziert das Risiko der Behinderungsprogression	✓	✓	Nicht bestimmbar	=
Verbessert die Behinderungsparameter EDSS MSFC	✓✓ ✓	✓ ✓✓	✓ ✓✓	Nicht bestimmbar
Verbessert die gesundheitsbezogene Lebensqualität	✓	✓	✓	Nicht bestimmbar
Bietet ein gutes Sicherheitsprofil	≈ (und Vermeidung grippeartiger Symptome)	≈ (und Vermeidung grippeartiger Symptome)	≈ (und Vermeidung grippeartiger Symptome)	✓
<p>✓✓: signifikant überlegen ✓: tendenziell überlegen =: kein Unterschied nachweisbar ≈: vergleichbar -: unterlegen a: nur overall p-Wert ermittelt über milde, mittelschwere und schwere Schübe b: für das Anwendungsgebiet a1 können diese Daten nur als Annäherung betrachtet werden, da keine Auswertung der relevanten Subpopulation möglich war</p>				

Insgesamt zeigen sich die Effekte in den relevanten Subpopulationen konsistent mit der Gesamtpopulation. Insbesondere zeigt sich eine Überlegenheit von Fingolimod hinsichtlich der Reduktion der jährlichen Schubrate und des Schubrisikos und damit der Erhöhung des schubfreien Patientenanteils. Die Gesamtzahl der Schübe und konsistent auch die Zahl schwerer Schübe oder Schübe, die einen Krankenhausaufenthalt beziehungsweise eine Steroidbehandlung erfordern, ist bei mit Fingolimod behandelten Patienten reduziert. Diese Schübe sind im Hinblick auf die Kosten, die für das Gesundheitssystem entstehen, besonders

relevant. Die Kosten, die durch einen hospitalisierungspflichtigen Schub entstehen, können bis um den Faktor 50 über jenen eines milden Schubes liegen wenn die Kosten für Neuro-Rehabilitationsmaßnahmen einbezogen werden (2). Selbst ohne diese Aufwendungen können die Kosten einen Wert von ungefähr 2.300 EUR pro Schub erreichen, bei Notwendigkeit einer Plasmapherese sogar bis zu rund 9.000 EUR. Eine Reduktion der Anzahl hospitalisierungspflichtiger Schübe ist daher besonders bedeutsam. Die Reduktion der Anzahl schwerer Schübe und hospitalisierungspflichtiger Schübe war in der Gesamtpopulation sowie den gemäß Label relevanten Subpopulationen zu sehen. Besonders auffällig ist, dass in der Subpopulation der Patienten mit schnell fortschreitender schwerer RRMS die Anzahl schwerer und hospitalisierungspflichtiger Schübe nicht nur konsistent mit der Gesamtzahl der Schübe sinkt, sondern der Anteil Schübe dieser Schubkategorie sogar prozentual deutlich reduziert wird. Auch wenn also die Schubrate in dieser Subpopulation nur mäßig und nicht signifikant gesenkt werden kann, so zeigt sich doch, dass die Schubschwere deutlich abnimmt und damit ein reduzierender Effekt auf die dadurch entstehenden Kosten zu erwarten ist.

Fingolimod reduziert außerdem das Risiko für eine Behinderungsprogression und verbessert die Behinderungsparameter. Die Lebensqualität kann konstant gehalten werden bzw. verbessert sich leicht. Unterschiede zwischen den relevanten Patientenpopulationen sind in Tabelle 4-82 dargestellt. Wie der Tabelle zu entnehmen ist, ist Fingolimod der zweckmäßigen Vergleichstherapie IFN- β -1a i. m. in nahezu allen Punkten signifikant oder zumindest tendenziell überlegen, speziell auch in den relevanten Anwendungsgebieten für diesen Vergleich (Anwendungsgebiete a2 und b). Fingolimod bietet ein gutes Sicherheitsprofil; grippeartige Symptome, wie sie bei der Gabe von IFN- β -1a i. m. beobachtet werden, haben unter Fingolimod keine Bedeutung. Ein Zusatznutzen von Fingolimod gegenüber IFN- β -1a i. m. kann daher festgestellt werden. Für den Vergleich mit Glatirameracetat stehen nicht zu allen genannten Endpunkten Vergleiche zur Verfügung, die Wahl der Endpunkte für den indirekten Vergleich durch die Publikationsangaben limitiert war. Insgesamt zeigt die obige Darstellung aber auch hier einen Zusatznutzen von Fingolimod im relevanten Anwendungsgebiet a1.

Der Zusatznutzen von Fingolimod gegenüber den zweckmäßigen Vergleichstherapien hinsichtlich patientenrelevanter Endpunkte wie der Reduktion der Schubrate und der Behinderungsprogression sowie das gut untersuchte Sicherheitsprofil befürworten Fingolimod zur Therapie der RRMS in den relevanten Patientenpopulationen. Das Sicherheits- und Verträglichkeitsprofil von Fingolimod ist gut charakterisiert und stellt insgesamt auch eine für den Patienten subjektiv wahrnehmbare Verbesserung gegenüber der zweckmäßigen Vergleichstherapie dar. Der Zusatznutzen hinsichtlich der Wirksamkeit und dem Ausbleiben von alltagseinschränkenden applikationsbedingten Nebenwirkungen kann daher die Therapiemotivation und die damit einhergehende Therapieadhärenz entscheidend verbessern. Die Ergebnisse zum medizinischen Zusatznutzen erlauben die Schlussfolgerung, dass Fingolimod eine langfristige Freiheit von schwerwiegenden Symptomen möglich macht. Fingolimod zeigt eine nachhaltige und bisher nicht erreichte Verbesserung des therapielevanten Nutzens in allen relevanten Patientenpopulationen. Der patientenrelevante Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie kann daher in allen relevanten

Patientenpopulationen als erheblich beschrieben werden und kann aufgrund der Studie der Klasse 1-Evidenz, die der Bewertung zugrunde liegt, als gesichert angesehen werden.

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 4-83: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
IFN non responder (compliant) (Anwendungsgebiet a1)	erheblich
IFN Non-Responder (non-compliant) (Anwendungsgebiet a2)	erheblich
Patienten mit schnell fortschreitender MS (Anwendungsgebiet b)	erheblich
a: siehe Modul 3 Abschnitt 3-2-1	

4.4.4 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens

Für Arzneimittel, die zur Behandlung eines seltenen Leidens zugelassen sind, müssen keine Nachweise zum medizinischen Nutzen und medizinischen Zusatznutzen vorgelegt werden (Abschnitt 4.3), solange der Umsatz des Arzneimittels mit der gesetzlichen Krankenversicherung zu Apothekenverkaufspreisen einschließlich Umsatzsteuer in den letzten 12 Monaten einen Betrag von 50 Millionen Euro nicht übersteigt.

Für diese Arzneimittel sind Angaben zu Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, im vorliegenden Abschnitt einzureichen. Dabei soll zunächst das Ausmaß des Zusatznutzens (in der Definition gemäß AM-NutzenV) begründet beschrieben werden. Unter Berücksichtigung des Ausmaßes des Zusatznutzens sollen Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, benannt werden. Weitere Abschnitte von Modul 4 können dabei nach Bedarf bearbeitet werden.

Beschreiben Sie begründet das Ausmaß des Zusatznutzens unter Verwendung der Kategorisierung gemäß AM-NutzenV. Berücksichtigen Sie bei den Aussagen zum Zusatznutzen nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen. Verweisen Sie, wo notwendig, auf weitere Abschnitte von Modul 4.

Nicht zutreffend.

Geben Sie auf Basis des Ausmaßes des Zusatznutzens in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

Tabelle 4-84: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Nicht zutreffend.	

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.

Für die in Modul 3 definierte zweckmäßige Vergleichstherapie Glatirameracetat in der Patientenpopulation der IFN Non-Responder (compliant) existieren keine direkten Vergleichsstudien zu Fingolimod. Aus diesem Grund war ein indirekter Vergleich nötig. Mit Fingolimod existieren RCT gegen Placebo und IFN- β -1a i. m. Für eine Netzwerkanalyse kommen daher alle RCT in Frage, die Glatirameracetat mit Placebo oder IFN- β -1a i. m. direkt vergleichen. Vergleiche zwischen den IFN- β -Präparaten untereinander und gegen Placebo existieren ebenso, daher kann eine Verknüpfung zu Glatirameracetat-Studien auch für Glatirameracetat-Vergleiche gegen ein anderes IFN- β -Präparat als IFN- β -1a i. m. erfolgen. Die Forderung war dabei, dass die Verbindung zu Fingolimod über maximal 2 Knoten erfolgte, da jeder zusätzliche Knoten den Standardfehler des indirekten Vergleichs erhöht. Der Vergleich wird, wie bei Hawkins et al. beschrieben, unzuverlässig, wenn mehr als ein paar wenige Knoten einbezogen werden (85). Daher wurde entschieden, für das Netzwerk nur Studien und Interventionen einzubeziehen, die Fingolimod mit höchstens 2 Zwischenstationen (Knoten) mit Glatirameracetat verbinden.

Die in den indirekten Vergleich eingeschlossenen Studien erfüllen diese Anforderungen. Für den indirekten Vergleich konnten allerdings keine Subpopulationen gebildet werden, da die in den dafür herangezogenen Publikationen Informationen dafür nicht ausreichten. Gemäß der vordefinierten Kriterien konnten keine Studien für einen indirekten Vergleich von Glatirameracetat mit Fingolimod identifiziert werden, die die IFN Non-Responder definieren. Entweder war eine Vortherapie von vornherein ausgeschlossen, oder die Vortherapien konnten keine IFN beinhalten, weil die Studien vor Einführung der IFN-Präparate rekrutiert hatten. Ein indirekter Vergleich mit Bezug auf die für Fingolimod relevanten Populationen ist daher nicht durchführbar. Die Suche wurde für diese Fragestellung daher auf alle erwachsenen RRMS-Patienten erweitert. Die verwendeten Publikationen sind zur Fragestellung die best-verfügbare Evidenz.

Die durchgeführte Netzwerkanalyse erfüllt den momentanen Stand der Wissenschaften hinsichtlich der Anforderungen an eine solche Analyse. Der Netzwerkvergleich ist daher aussagekräftig.

4.5.2 Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nichtrandomisierte 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 2005¹⁴, Molenberghs 2010¹⁵). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006¹⁶) 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 2006¹⁷) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

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

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

Sofern Sie im Dossier Ergebnisse zu Surrogatendpunkten eingereicht haben, benennen Sie die Gründe für die Verwendung von Surrogatendpunkten. Beschreiben Sie, ob und warum die

¹⁴ Burzykowski T (Ed.: *The evaluation of surrogate endpoints*. New York: Springer; 2005.

¹⁵ 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.

¹⁶ Burzykowski T, Buyse M. *Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation*. *Pharm Stat* 2006; 5(3): 173-186.

¹⁷ Weir CJ, Walley RJ. *Statistical evaluation of biomarkers as surrogate endpoints: a literature review*. *Stat Med* 2006; 25(2): 183-203.

verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Nicht zutreffend.

4.6 Liste der eingeschlossenen Studien

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Datenquellen auf.

FREEDOMS – FTY720D2301; Studienbericht, Publikation, Registereinträge (22, 25,26, 29, 30)

TRANSFORMS – FTY720D2302; Studienbericht, Publikation, Registereinträge (23, 27-29, 31)

EVIDENCE, Publikationen, Registereinträge (45, 46, 75-79)

INCOMIN, Publikation (43)

MSCRG, Publikation (62-66)

Johnson, Publikation (50, 51, 67-69)

The IFNB MS Study Group, Publikation (39, 70-72)

PRISMS, Publikation (40, 73)

BEYOND, Publikation, Registereinträge (80, 52-54)

BECOME, Publikation, Registereinträge (81, 55, 56)

REGARD, Publikation, Registereinträge (82, 57, 58)

Etemadifar, Publikation (83)

Bornstein, Publikation (74)

Indirekter Vergleich (Netzwerkanalyse), RTI Bericht (24)

4.7 Referenzliste

Benennen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregister-einträge), die Sie in Modul 4 angegeben haben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard).

- (1) Gemeinsamer Bundesausschuss. Niederschrift zum Beratungsgespräch gemäß § 8 Abs. 1 AM-NutzenV. Beratungsanforderung 2011-04-15-D-004. Datum des Gesprächs 01.07.2011. Erhalt der Niederschrift 13.07.2011.
- (2) O'Brien J, Ward A, Patrick A, Caro J. Cost of managing an episode of relapse in multiple sclerosis in the United States. BMC Health Services Research 2003 Sep 2;3(1):17.
- (3) DiMarco JP, O'Connor P, Cohen JA, Francis G, Collins W, Zhang-Auberson L, et al. First-dose effect of fingolimod: pooled safety data from two phase 3 studies (TRANSFORMS and FREEDOMS). Mult Scler 2010.
- (4) Schmouder R, Serra D, Wang Y, Kovarik JM, DiMarco J, Hunt TL, et al. FTY720: placebo-controlled study of the effect on cardiac rate and rhythm in healthy subjects. J Clin Pharmacol 2006;46(8):895-904.
- (5) Biogen Idec Limited. Fachinformation Avonex[®] 30 Mikrogramm/0,5ml. www fachinfo de 2010 March [cited 2011 Mar 11];Available from: URL: www.fachinfo.de
- (6) Merck. Fachinformation Rebif[®] 22 Mikrogramm und 44 Mikrogramm. www fachinfo de 2010 August [cited 2011 Mar 23];Available from: URL: www.fachinfo.de
- (7) Bayer Schering Pharma. Fachinformation Betaferon[®] 250 Mikrogramm/ml. www fachinfo de 2011 January [cited 2011 Mar 23];Available from: URL: www.fachinfo.de
- (8) Novartis Pharma. Fachinformation Extavia[®]. www fachinfo de 2009 November [cited 2011 Mar 23];
- (9) Teva, Sanofi Aventis. Fachinformation Copaxone[®] 20mg/ml. www fachinfo de 2009 March [cited 2011 Mar 23];Available from: URL: www.fachinfo.de
- (10) European Medicines Agency. Guideline on Clinical Investigation of Medicinal Products for the Treatment of Multiple Sclerosis. 1-6-2007.
- (11) European Union. COUNCIL DIRECTIVE on the approximation of the laws of Member States relating to analytical, pharmaco-toxicological and clinical standards and protocols in respect of the testing of proprietary medicinal products (75/318/EEC). 20-5-1975.
- (12) International Conference of Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human use. The Extent of Population Exposure to Assess Clinical Safety for Drugs intended for Long-Term Treatment of Non-Life-Threatening Conditions - E1. 1994 Oct 27.

- (13) International Conference of Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human use. Dose-Response Information to Support Drug Registration - E4. 1994 Mar 10.
- (14) International Conference of Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human use. Statistical Principles for Clinical Trials - E9. 1998 Feb 5.
- (15) International Conference of Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human use. Choice of Control Group and Related Issues in Clinical Trials - E10. 2000 Jul 20.
- (16) IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden Version 3.0. 2008 May 27.
- (17) Kurtzke JF. A New Scale for Evaluating Disability in Multiple Sclerosis. *Neurology* 1955 Aug 1;5(8):580-3.
- (18) Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983 Nov;33(11):1444-52.
- (19) Cutter GR, Baier ML, Rudick RA, Cookfair DL, Fischer JS, Petkau J, et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. *Brain* 1999 May;122 (Pt 5):871-82.
- (20) Fisk JD, Brown MG, Sketris IS, Metz LM, Murray TJ, Stadnyk KJ. A comparison of health utility measures for the evaluation of multiple sclerosis treatments. *J Neurol Neurosurg Psychiatry* 2005 Jan;76(1):58-63.
- (21) Greiner W, Claes C, Busschbach JJ, von der Schulenburg JM. Validating the EQ-5D with time trade off for the German population. *Eur J Health Econ* 2005 Jun;6(2):124-30.
- (22) Novartis. FTY720D2301 (FREEDOMS) - Studienbericht. 2009.
- (23) Novartis. FTY720D2302 (TRANSFORMS) - Studienbericht. 2008.
- (24) RTI Health Solutions. Systematic Review and Meta-analysis of selected therapies for relapsing-remitting multiple sclerosis. 2011.
- (25) Novartis. NCT00289978, ClinicalTrials.gov Registereintrag. <http://www.clinicaltrial.gov> . 22-2-2011. 15-8-2011.
- (26) Novartis. NCT00289978, ICTRP Registereintrag FREEDOMS. <http://apps.who.int/trialsearch/default.aspx> . 3-6-2010. 9-9-2011.
- (27) Novartis. NCT00340834, ClinicalTrials.gov Registereintrag. <http://www.clinicaltrial.gov> . 20-4-2011. 15-8-2011.

- (28) Novartis. NCT00340834, ICTRP Registereintrag. <http://apps.who.int/trialsearch/default.aspx> . 3-6-2010. 9-9-2011.
- (29) Novartis. Zusatzauswertungen TRANSFORMS und FREEDOMS. 2011 Sep 16.
- (30) Kappos L, Radue EW, O'Connor P, Polman C, Hohlfeld R, Calabresi P, et al. A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *N Engl J Med* 2010 Feb 4;362(5):387-401.
- (31) Cohen JA, Barkhof F, Comi G, Hartung HP, Khatri BO, Montalban X, et al. Oral fingolimod or intramuscular interferon for relapsing multiple sclerosis. *N Engl J Med* 2010 Feb 4;362(5):402-15.
- (32) Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Ann Neurol* 2005 Dec;58(6):840-6.
- (33) Bundesministerium für Gesundheit. ICD-10-GM Version 2011, Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme. 2011. Deutsches Institut für Medizinische Dokumentation und Information, DIMDI. 24-9-2010.
- (34) DRG-Browser, Report-Browser für die Kalkulationsergebnisse bei Versorgung durch Hauptabteilungen (HA) [computer program]. Version V2009/2011 2011.
- (35) Institut für das Entgeltsystem im Krankenhaus InEK GmbH. Fallpauschalen-Katalog 2011. 2011.
- (36) GKV-Spitzenverband, Verband der privaten Krankenversicherung e.V., Deutsche Krankenhausgesellschaft. Vereinbarung gemäß § 10 Abs. 9 KHEntgG für den Vereinbarungszeitraum 2011. 2010.
- (37) Wiendl H, Kieseier BC. Eskalationstherapie des Schubes. In: Brandt T, Hohlfeld R, Noth J, Reichmann H, editors. *Multiple Sklerose - Klinik, Diagnostik und Therapie*. 1. Auflage ed. Kohlhammer; 2011. p. 161-2.
- (38) National Cancer Institute. Probability of Breast Cancer in American Women. 2009 [cited 2009 Nov 3]; Available from: URL: <http://www.cancer.gov/cancertopics/factsheet/Detection/probability-breast-cancer/print?page=&keyword=>
- (39) The IFNB Multiple Sclerosis Study Group. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. I. Clinical results of a multicenter, randomized, double-blind, placebo-controlled trial. *Neurology* 1993 Apr;43(4):655-61.
- (40) PRISMS (Prevention of Relapses and Disability by Interferon beta-1a Subcutaneously in Multiple Sclerosis) Study Group. Randomised double-blind placebo-controlled study of interferon beta-1a in relapsing/remitting multiple sclerosis. *Lancet* 1998 Nov 7;352(9139):1498-504.

- (41) Jacobs LD, Cookfair DL, Rudick RA, Herndon RM, Richert JR, Salazar AM, et al. Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. The Multiple Sclerosis Collaborative Research Group (MSCRG). *Ann Neurol* 1996 Mar;39(3):285-94.
- (42) Mohr DC, Goodkin DE, Masuoka L, Dick LP, Russo D, Eckhardt J, et al. Treatment adherence and patient retention in the first year of a Phase-III clinical trial for the treatment of multiple sclerosis. *Mult Scler* 1999 Jun;5(3):192-7.
- (43) Durelli L, Verdun E, Barbero P, Bergui M, Versino E, Ghezzi A, et al. Every-other-day interferon beta-1b versus once-weekly interferon beta-1a for multiple sclerosis: results of a 2-year prospective randomised multicentre study (INCOMIN). *Lancet* 2002;359(9316):1453-60.
- (44) Datamonitor. Stakeholder Insight: Multiple Sclerosis. Disease-modifying efficacy and side effects guide treatment choice. 2006 [cited 2008 Nov 10];2008 Available from: URL:
[http://www.datamonitor.com/kc/pharma/iProduct/toc/index.asp?R=DMHC2256&ref=findInfo/index.asp?search=multiple+sclerosis+AND+treatment+guidelines\\$;26;\\$x=5\\$;26;\\$y=7\\$;26;\\$N=%2D112131+337&Npreserve=337&defaultSearch=multiple%20sclerosis%20AND%20treatment%20guidelines](http://www.datamonitor.com/kc/pharma/iProduct/toc/index.asp?R=DMHC2256&ref=findInfo/index.asp?search=multiple+sclerosis+AND+treatment+guidelines$;26;$x=5$;26;$y=7$;26;$N=%2D112131+337&Npreserve=337&defaultSearch=multiple%20sclerosis%20AND%20treatment%20guidelines)
- (45) EMD Serono. NCT00292266, ClinicalTrials.gov Registereintrag. <http://www.clinicaltrial.gov> . 21-3-2009. 25-8-2011.
- (46) EMD Serono. NCT00292266, ICTRP Registereintrag. <http://apps.who.int/trialsearch/default.aspx> . 3-6-2010. 9-9-2011.
- (47) Bayer HealthCare. NCT00206648, Clinicaltrials.gov Registereintrag. 7-11-2008. 12-9-2011.
- (48) Bayer HealthCare. NCT00206648, ICTRP Registereintrag. 3-6-2010. 12-9-2011.
- (49) Bayer HealthCare. NCT00206648, clinicalstudyresults.org Registereintrag und Ergebnis. 2007. 12-9-2011.
- (50) National Center for Research Resources, University of Maryland. NCT00004814, ClinicalTrials.gov Registereintrag. 21-3-2009. 12-9-0011.
- (51) National Center for Research Resources, University of Maryland. NCT00004814, ICTRP Registereintrag. 24-2-2000. 12-9-0011.
- (52) Bayer HealthCare. NCT00099502, ClinicalTrials.gov Registereintrag. <http://www.clinicaltrial.gov> . 18-12-2008. 25-8-2011.
- (53) Bayer HealthCare. NCT00099502, ICTRP Registereintrag. <http://apps.who.int/trialsearch/default.aspx> . 3-6-2010. 9-9-2011.
- (54) Bayer HealthCare. NCT00099502, clinicalstudyresults.org Registereintrag und Ergebnis. 31-7-2008. 12-9-2011.

- (55) University of Medicine and Dentistry New Jersey. NCT00176592, ClinicalTrials.gov Registereintrag. <http://www.clinicaltrial.gov> . 6-5-2010. 25-8-2011.
- (56) University of Medicine and Dentistry New Jersey. NCT00176592, ICTRP Registereintrag. <http://apps.who.int/trialsearch/default.aspx> . 3-6-2010. 9-9-2011.
- (57) EMD Serono. NCT00078338, ClinicalTrials.gov Registereintrag. <http://www.clinicaltrial.gov> . 19-4-2011. 25-8-2011.
- (58) EMD Serono. NCT00078338, ICTRP Registereintrag. <http://apps.who.int/trialsearch/default.aspx> . 26-4-2011. 9-9-2011.
- (59) Teva Pharmaceutical Industries. NCT00202995, ClinicalTrials.gov Registereintrag. 14-10-2010. 12-9-2011.
- (60) Teva Pharmaceutical Industries. NCT00202995, ICTRP Registereintrag. 13-9-2005. 12-9-2011.
- (61) EMD Serono, Pfizer. NCT00101959, ClinicalTrials.gov Registereintrag. 21-3-2009. 12-9-2011.
- (62) Miller DM, Weinstock-Guttman B, Bourdette D, You X, Foulds P, Rudick RA. Change in quality of life in patients with relapsing-remitting multiple sclerosis over 2 years in relation to other clinical parameters: Results from a trial of intramuscular interferon beta-1a. *Multiple Sclerosis*
- (63) Rudick RA, Lee JC, Cutter GR, Miller DM, Bourdette D, Weinstock-Guttman B, et al. Disability progression in a clinical trial of relapsing-remitting multiple sclerosis eight-year follow-up. *Archives of Neurology* 2010 Nov;67(11):1329-35.
- (64) Rudick RA, Goodkin DE, Jacobs LD, Cookfair DL, Herndon RM, Richert JR, et al. Impact of interferon beta-1a on neurologic disability in relapsing multiple sclerosis. *Neurology* 2001 Dec 26;57(12 SUPPL. 5):S25-S30.
- (65) Rudick RA, Goodkin DE, Jacobs LD, Cookfair DL, Herndon RM, Richert JR, et al. Impact of interferon beta-1a on neurologic disability in relapsing multiple sclerosis. *Neurology* 1997 Aug;49(2):358-63.
- (66) Jacobs LD, Cookfair DL, Rudick RA, Herndon RM, Richert JR, Salazar AM, et al. A phase III trial of intramuscular recombinant interferon beta as treatment for exacerbating-remitting multiple sclerosis: design and conduct of study and baseline characteristics of patients. *Multiple Sclerosis Collaborative Research Group (MSCRG). Multiple sclerosis (Houndmills, Basingstoke, England)* 1995 Jun;1(2):118-35.
- (67) Johnson KP, Brooks BR, Cohen JA, Ford CC, Goldstein J, Lisak RP, et al. Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: results of a phase III multicenter, double-blind placebo-controlled trial. *Neurology* 1995 Jul;45(7):1268-76.

- (68) Johnson KP, Brooks BR, Cohen JA, Ford CC, Goldstein J, Lisak RP, et al. Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: Results of a phase III multicenter, double-blind, placebo-controlled trial. *Neurology*
- (69) Liu C, Blumhardt LD. Benefits of glatiramer acetate on disability in relapsing-remitting multiple sclerosis: An analysis by area under disability/time curves. *Journal of the Neurological Sciences*
- (70) Interferon beta-1b in the treatment of multiple sclerosis: final outcome of the randomized controlled trial. The IFNB Multiple Sclerosis Study Group and The University of British Columbia MS/MRI Analysis Group. *Neurology*
- (71) Duquette P, Girard M, Despault L, Dubois R, Knobler RL, Lublin FD, et al. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. I. Clinical results of a multicenter, randomized, double-blind, placebo-controlled trial. *Neurology* 1993;43(4 I):655-61.
- (72) Duquette P, Girard M, Despault L, Dubois B, Knobler RL, Lublin FD, et al. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. I. Clinical results of a multicenter, randomized, double-blind, placebo-controlled trial. *Neurology* 2001 Dec 26;57(12 SUPPL. 5):S3-S9.
- (73) Liu C, Blumhardt LD. Randomised, double blind, placebo controlled study of interferon beta-1a in relapsing-remitting multiple sclerosis analysed by area under disability/time curves. *Journal of Neurology Neurosurgery and Psychiatry* 1999 Oct;67(4):451-6.
- (74) Bornstein MB, Miller A, Slagle S, Weitzman M, Crystal H, Drexler E, et al. A pilot trial of Cop 1 in exacerbating-remitting multiple sclerosis. *N Engl J Med* 1987 Aug 13;317(7):408-14.
- (75) Panitch H, Goodin DS, Francis G, Chang P, Coyle PK, O'Connor P, et al. Randomized, comparative study of interferon beta-1a treatment regimens in MS: The EVIDENCE Trial. *Neurology* 2002 Nov 26;59(10):1496-506.
- (76) Schwid SR, Panitch HS. Full results of the Evidence of Interferon Dose-Response-European North American Comparative Efficacy (EVIDENCE) study: A multicenter, randomized, assessor-blinded comparison of low-dose weekly versus high-dose, high-frequency interferon beta-1a for relapsing multiple sclerosis. *Clinical Therapeutics* 2007 Sep;29(9):2031-48.
- (77) Schwid SR, Thorpe J, Sharief M, Sandberg-Wollheim M, Rammohan K, Wendt J, et al. Enhanced Benefit of Increasing Interferon Beta-1a Dose and Frequency in Relapsing Multiple Sclerosis: The EVIDENCE Study. *Archives of Neurology* 2005 May 1;62(5):785-92.
- (78) Sandberg-Wollheim M, Bever C, Carter J, Fikris M, Hurwitz B, Lapierre Y, et al. Comparative tolerance of IFN beta-1a regimens in patients with relapsing multiple sclerosis. *Journal of Neurology* 2005 Jan 1;252(1):8-13.

- (79) Panitch H, Goodin D, Francis G, Chang P, Coyle P, O'Connor P, et al. Benefits of high-dose, high-frequency interferon beta-1a in relapsing-remitting multiple sclerosis are sustained to 16 months: Final comparative results of the EVIDENCE trial. *Journal of the Neurological Sciences* 2005 Dec 15;239(1):67-74.
- (80) O'Connor P, Filippi M, Arnason B, Comi G, Cook S, Goodin D, et al. 250 µg or 500 µg interferon beta-1b versus 20 mg glatiramer acetate in relapsing-remitting multiple sclerosis: a prospective, randomised, multicentre study. *Lancet Neurol* 2009 Sep 2;8:889-97.
- (81) Cadavid D, Wolansky LJ, Skurnick J, Lincoln J, Cheriyan J, Szczepanowski K, et al. Efficacy of treatment of MS with IFNbeta-1b or glatiramer acetate by monthly brain MRI in the BECOME study. *Neurology* 2009 Jun 9;72(23):1976-83.
- (82) Mikol DD, Barkhof F, Chang P, Coyle PK, Jeffery DR, Schwid SR, et al. Comparison of subcutaneous interferon beta-1a with glatiramer acetate in patients with relapsing multiple sclerosis (the REbif vs Glatiramer Acetate in Relapsing MS Disease [REGARD] study): a multicentre, randomised, parallel, open-label trial. *Lancet Neurol* 2008 Oct;7(10):903-14.
- (83) Etemadifar M, Janghorbani M, Shaygannejad V. Comparison of Betaferon, Avonex, and Rebif in treatment of relapsing-remitting multiple sclerosis. *Acta Neurol Scand* 2006 May;113(5):283-7.
- (84) Klawiter EC, Cross AH, Naismith RT. The present efficacy of multiple sclerosis therapeutics: Is the new 66% just the old 33%? *Neurology* 2009 Sep 22;73(12):984-90.
- (85) Hawkins N, Scott SA, Woods B. How far do you go? Efficient searching for indirect evidence. *Med Decis Making* 2009;29(3):273-81.
- (86) Novartis. NCT00731692, ClinicalTrials.gov Registereintrag. <http://www.clinicaltrial.gov> . 12-10-2010. 15-8-2011.
- (87) Novartis. NCT00731692, ICTRP Registereintrag. <http://apps.who.int/trialsearch/default.aspx> . 6-9-2011. 9-9-2011.
- (88) Novartis. NCT00355134, ClinicalTrials.gov Registereintrag. <http://www.clinicaltrial.gov> . 30-9-2010. 15-8-2011.
- (89) Novartis. NCT00355134, ICTRP Registereintrag. <http://apps.who.int/trialsearch/default.aspx> . 5-10-2010. 9-9-2011.
- (90) Novartis. NCT01216072, ClinicalTrials.gov Registereintrag. <http://www.clinicaltrial.gov> . 2-5-2011. 15-8-2011.
- (91) Novartis. NCT01216072, ICTRP Registereintrag. <http://apps.who.int/trialsearch/default.aspx> . 10-5-2011. 9-9-2011.
- (92) Novartis. NCT01333501, ClinicalTrials.gov Registereintrag. 18-5-2011. 12-9-2011.

- (93) Novartis. NCT01333501, ICTRP Registereintrag. 24-5-2011. 12-9-2011.
- (94) Novartis. NCT01317004, ClinicalTrials.gov Registereintrag. 31-5-2011. 12-9-2011.
- (95) Novartis. NCT01317004, ICTRP Registereintrag. 7-6-2011. 12-9-2011.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die in Abschnitt 4.2.3.2 beschriebene bibliografische Literaturrecherche an, und zwar einzeln für jede Datenbank. 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. Geben Sie zusätzlich an, für welchen Abschnitt von Modul 4 die Suche durchgeführt wurde.

Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	08.12.2010	
Zeitsegment	1980 to 2010 week 50	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ¹⁸] – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	Meglitinide/	848
2	Nateglinide/	1686
3	Repaglinide/	2118
4	(glinid* or meglitinid* or nateglinid* or repaglinid*).ab,ti.	1069
5	(starlix or novonorm or novo norm or prandin).ab,ti.	32
6	(105816-04-4 or 135062-02-1).rn.	2854
7	or/1-6	3467
8	Diabetes mellitus/	224164
9	Non Insulin dependent Diabetes mellitus/	91081
10	(diabet* or niddm or t2dm).ab,ti.	379777
11	or/8-10	454517
12	(random* or double-blind*).tw.	650136
13	Placebo*.mp.	243550
14	or/12-13	773621
15	and/7,11,14	719

¹⁸ 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.

Fingolimod, Abschnitt 4.3.1

Datenbank	MEDLINE /MEDLINE In-Process & Other Non-Indexed Citations	
Suchoberfläche	Ovid	
Datum der Suche	25.07.2011	
Zeitsegment	1948 to July Week 2 2011 / July 22 2011	
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	(fingolimod or gilenia or gilenya or fty720* or fty 720* or pkf117-812 or pkf117812).mp.	964
2	(162359-56-0 or 162359-55-9).rn.	744
3	or/1-2	964
4	exp multiple sclerosis/	37350
5	(multiple sclerosis or RRMS or MS).ti,ab.	178422
6	or/4-5	184038
7	randomized controlled trial.pt.	311981
8	(randomized or placebo).mp.	496809
9	or/7-8	496809
10	and/3,6,9	23

Datenbank	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	25.07.2011	
Zeitsegment	1980 to 2011 Week 29	
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.	
1	Exp fingolimod/	1947
2	(fingolimod or gilenia or gilenya or fty720* or fty 720* or pkf117-812 or pkf117812).mp.	2019
3	(162359-56-0 or 162359-55-9).rn.	1815
4	or/1-3	2019
5	exp multiple sclerosis/	56485
6	(multiple sclerosis or RRMS or MS).ti,ab.	209884
7	or/5-6	222355
8	(random* or double-blind*).tw.	670720
9	placebo*.mp.	252109
10	or/8-9	800472
11	and/4,7,10	249

Datenbank	COCHRANE Database of Systematic Reviews	
Suchoberfläche	Ovid	
Datum der Suche	25.07.2011	
Zeitsegment	2005 to July 2011	
Suchfilter	---	

#	Suchbegriffe	Ergebnis
1	(fingolimod or gilenia or gilenya or fty720* or fty 720* or pkf117-812 or pkf117812).af.	2
2	(multiple sclerosis or RRMS or MS).af.	2458
3	and/1-2	1

Datenbank	COCHRANE Central Register of Controlled Trials	
Suchoberfläche	Ovid	
Datum der Suche	25.07.2011	
Zeitsegment	3rd Quarter 2011	
Suchfilter	--	
#	Suchbegriffe	Ergebnis
1	(fingolimod or gilenia or gilenya or fty720* or fty 720* or pkf117-812 or pkf117812).af.	75
2	(multiple sclerosis or RRMS or MS).af.	13165
3	and/1-2	27

IFN- β -Präparate, Abschnitt 4.3.2.1IFN- β -1a vs. IFN- β -1b:

Datenbank	MEDLINE /MEDLINE In-Process & Other Non-Indexed Citations	
Suchoberfläche	Ovid	
Datum der Suche	28.07.2011	
Zeitsegment	1948 to July Week 3 2011 / July 27 2011	
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	(betaseron or betaferon or extavia or nvf233 or nvf 233 or sh4579a or sh 4579a or shy03967a or shy03967a).af.	207
2	145155-23-3.rn.	762
3	((interferon* or ifn* or peginterferon* or pegifn*) adj3 (1b or lb or 1 b or l b or beta1b or betalb or beta1 b or betal b or ser or serin*)).af.	1721
4	or/1-3	1760
5	(rebif or avonex or bioferon or cinnovex or bg9418 or bg 9418).af.	295
6	145258-61-3.rn.	1082
7	((interferon* or ifn* or peginterferon* or pegifn*) adj3 (1a or la or 1 a or l a or beta1a or betala or beta1 a or betal)).af.	1627
8	or/5-7	1666
9	And/4,8	715
10	exp multiple sclerosis/	37395
11	(multiple sclerosis or RRMS or MS).ti,ab.	178626
12	or/10,11	184246
13	(compar* or head or versus or vs).mp.	4157632
14	comparative study.pt.	1538448
15	exp comparative effectiveness research/	464
16	(randomized or placebo).mp.	497348
17	randomized controlled trial.pt.	312372
18	Or/13-17	4349384
19	And/9,12,18	312

Datenbank	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	28.07.2011	
Zeitsegment	1980 to 2011 Week 29	
Suchfilter	---	
1	exp *beta1a interferon/	1220
2	(rebif or avonex or bioferon or cinnovex or bg9418 or bg 9418).ti,ab.	449
3	((interferon* or ifn* or peginterferon* or pegifn*) adj3 (1a or la or 1 a or l a or beta1a or betala or beta1 a or	1513

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

	betal)).ti,ab.	
4	or/1-3	2140
5	exp *interferon beta serine/	1117
6	(betaseron or betaferon or extavia or nvf233 or nvf 233 or sh4579a or sh 4579a or shy03967a or shy03967a).ti,ab.	332
7	((interferon* or ifn* or peginterferon* or pegifn*) adj3 (1b or lb or 1 b or l b or beta1b or betalb or beta1 b or betal b or ser or serin*)).ti,ab.	1418
8	or/5-7	2053
9	(compar* or head or versus or vs).ti,ab.	3716683
10	exp comparative study/	856618
11	or/9,10	4120758
12	and/4,8,11	307
13	exp interferon beta serine/cm	346
14	exp beta1a interferon/cm	501
15	and/13-14	266
16	exp *recombinant interferon/	13816
17	exp *beta interferon/	5958
18	exp *recombinant beta interferon/	420
19	or/16-18	19343
20	and/15,19	161
21	or/12,20	390

Datenbank	COCHRANE Database of Systematic Reviews	
Suchoberfläche	Ovid	
Datum der Suche	28.07.2011	
Zeitsegment	2005 to July 2011	
Suchfilter	---	
#	Suchbegriffe	Ergebnis
1	(rebif or avonex or bioferon or cinnovex or bg9418 or bg 9418).mp.	8
2	((interferon* or ifn* or peginterferon* or pegifn*) adj3 (1a or la or 1 a or l a or beta1a or betala or beta1 a or betal)).mp.	43
3	or/1-2	44
4	(betaseron or betaferon or extavia or nvf233 or nvf 233 or sh4579a or sh 4579a or shy03967a or shy03967a).mp.	5
5	((interferon* or ifn* or peginterferon* or pegifn*) adj3 (1b or lb or 1 b or l b or beta1b or betalb or beta1 b or betal b or ser or serin*)).mp.	10
6	or/4-5	11
7	and/3,6	9
8	Multiple sclerosis	146
9	and/7-8	9

Datenbank	COCHRANE Central Register of Controlled Trials
-----------	------------------------------------------------

Suchoberfläche	Ovid	
Datum der Suche	28.07.2011	
Zeitsegment	3rd Quarter 2011	
Suchfilter	--	
#	Suchbegriffe	Ergebnis
1	(rebif or avonex or bioferon or cinnovex or bg9418 or bg 9418).mp	122
2	((interferon* or ifn* or peginterferon* or pegifn*) adj3 (1a or la or 1 a or l a or beta1a or betala or beta1 a or betal)).mp.	1261
3	or/1-2	1292
4	(betaseron or betaferon or extavia or nvf233 or nvf 233 or sh4579a or sh 4579a or shy03967a or shy03967a).mp.	49
5	((interferon* or ifn* or peginterferon* or pegifn*) adj3 (1b or lb or 1 b or l b or beta1b or betalb or beta1 b or betal b or ser or serin*)).mp.	272
6	or/4-5	295
7	and/3,6	69
8	Multiple sclerosis.mp.	2539
9	and/7-8	46
10	(compar* or versus or vs or head).mp.	318792
11	and/9-10	33

IFN- β -1a i. m. vs. IFN- β -1a s. c.:

Datenbank	MEDLINE /MEDLINE In-Process & Other Non-Indexed Citations	
Suchoberfläche	Ovid	
Datum der Suche	13.09.2011	
Zeitsegment	1948 to Sep Week 1 2011	
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	(sc or (s adj c) or subcutan* or subkutan*).mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier]	171824
2	(im or (i adj m) or intramusc* or (intra adj musc*) or intramusk* or (intra adj musk*)).mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier]	135693
3	145258-61-3.rn.	1099
4	((interferon* or ifn* or peginterferon* or pegifn*) adj3 (1a or la or ("1" adj a) or (l adj a) or beta1a or betala or (beta1 adj a) or betal)).af.	1644
5	rebif.af. ()	172
6	(avonex or cinnovex or bg9418 or (bg adj "9418")).af.	221
7	((3 or 4) and 1 and 2) or (6 and 1) or (5 and 2)	138
8	exp multiple sclerosis/	38216
9	(multiple sclerosis or RRMS or MS).ti,ab.	182593
10	8 or 9	188304
11	(compar* or head or versus or vs).mp.	4223502
12	comparative study.pt.	1554945
13	exp comparative effectiveness research/	527
14	(randomized or placebo).mp.	506055
15	randomized controlled trial.pt.	317076
16	or/11-15	4418843
17	and/7,10,16	100

Datenbank	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	13.09.2011	
Zeitsegment	1980 to 2011 Week 36	
Suchfilter	---	
1	(sc or (s adj c) or subcutan* or subkutan*).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]	257915
2	(im or (i adj m) or intramusc* or (intra adj musc*) or intramusk* or (intra adj musk*)).mp. [mp=title, abstract, subject headings, heading word, drug trade name,	206090

	original title, device manufacturer, drug manufacturer, device trade name, keyword]	
3	((interferon* or ifn* or peginterferon* or pegifn*) adj3 (1a or 1a or ("1" adj a) or (l adj a) or beta1a or betala or (beta1 adj a) or betal)).af.	4489
4	rebif.af.	1582
5	(avonex or cinnovex or bg9418 or (bg adj "9418")).af.	1942
6	(3 and 1 and 2) or (5 and 1) or (4 and 2)	310
7	(compar* or head or versus or vs).ti,ab.	3955336
8	exp comparative study/	916460
9	7 or 8	4399597
10	and/6,9	179
11	exp beta1a interferon/cm	515
12	exp *recombinant interferon/	13950
13	exp *beta interferon/	6015
14	exp *recombinant beta interferon/	420
15	or/12-14	19533
16	and/1-2,11,15	49
17	or/10,16	190

Datenbank	COCHRANE Database of Systematic Reviews	
Suchoberfläche	Ovid	
Datum der Suche	13.09.2011	
Zeitsegment	2005 to August 2011	
Suchfilter	---	
#	Suchbegriffe	Ergebnis
1	(sc or (s adj c) or subcutan* or subkutan*).mp. [mp=title, abstract, full text, keywords, caption text]	633
2	(im or (i adj m) or intramusc* or (intra adj musc*) or intramus* or (intra adj musk*)).mp. [mp=title, abstract, full text, keywords, caption text]	598
3	((interferon* or ifn* or peginterferon* or pegifn*) adj3 (1a or 1a or ("1" adj a) or (l adj a) or beta1a or betala or (beta1 adj a) or betal)).af.	54
4	rebif.af.	10
5	(avonex or cinnovex or bg9418 or (bg adj "9418")).af.	9
6	(3 and 1 and 2) or (5 and 1) or (4 and 2)	13
7	multiple sclerosis.mp. [mp=title, short title, abstract, full text, keywords, caption text]	146
8	and/6-7	8

Datenbank	COCHRANE Central Register of Controlled Trials	
Suchoberfläche	Ovid	
Datum der Suche	13.09.2011	
Zeitsegment	3rd Quarter 2011	
Suchfilter	--	
#		
1	(sc or (s adj c) or subcutan* or subkutan*).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]	10486
2	(im or (i adj m) or intramusc* or (intra adj musc*) or intramusk* or (intra adj musk*)).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]	11070
3	((interferon* or ifn* or peginterferon* or pegifn*) adj3 (1a or la or ("1" adj a) or (l adj a) or beta1a or betala or (beta1 adj a) or betal)).af.	1263
4	rebif.af.	66
5	(avonex or cinnovex or bg9418 or (bg adj "9418")).af.	72
6	(3 and 1 and 2) or (5 and 1) or (4 and 2)	26
7	multiple sclerosis.mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]	2539
8	and/6-7	15
9	(compar* or versus or vs or head).mp.	318792
10	and/8-9	14

IFN- β -Präparate vs. Placebo:

Datenbank	MEDLINE /MEDLINE In-Process & Other Non-Indexed Citations	
Suchoberfläche	Ovid	
Datum der Suche	28.07.2011	
Zeitsegment	1948 to July Week 3 2011 / July 27 2011	
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	interferon type I/ or Interferon Type I, Recombinant/ or Interferon-beta/ or interferons/ad, tu	22312
2	(extavia or betaferon or betaseron or rebif or avonex or interferonbeta* or ifnbeta* or ((peginterferon* or pegifn* or interferon* or ifn*) adj3 (beta* or 1a or 1b or la or lb or "1 a" or "1 b" or "l a" or "l b"))).ti,ab.	12941
3	or/1-2	27870
4	exp multiple sclerosis/	37395
5	(multiple sclerosis or RRMS or MS).ti,ab.	178626
6	or/4-5	184246
7	randomized controlled trial.pt.	312372
8	(randomized or placebo).mp.	497348
9	or/7-8	497348
10	and/3,6,9	557
11	placebo*.mp.	148336
12	and/10-11	301

Datenbank	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	28.07.2011	
Zeitsegment	1980 to 2011 Week 29	
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.	
1	*recombinant interferon/ or *recombinant beta interferon/ or *beta1a interferon/ or *interferon beta serine/ or *beta interferon/ or *beta1 interferon/	9002
2	(extavia or betaferon or betaseron or rebif or avonex or interferonbeta* or ifnbeta* or ((peginterferon* or pegifn* or interferon* or ifn*) adj3 (beta* or 1a or 1b or la or lb or "1 a" or "1 b" or "l a" or "l b"))).ti,ab.	15238
3	or/1-2	17966
4	exp multiple sclerosis/	56485
5	(multiple sclerosis or RRMS or MS).ti,ab.	209884
6	or/4-5	222355
7	(random* or double-blind*).tw.	670720
8	placebo*.mp.	252109
9	or/7-8	800472

10	and/3,6,9	934
11	placebo*.mp.	252109
12	and/10-11	704

Datenbank	COCHRANE Database of Systematic Reviews	
Suchoberfläche	Ovid	
Datum der Suche	28.07.2011	
Zeitsegment	2005 to July 2011	
Suchfilter	---	
#	Suchbegriffe	Ergebnis
1	(extavia or betaferon or betaseron or rebif or avonex or interferonbeta* or ifnbeta* or ((peginterferon* or pegifn* or interferon* or ifn*) adj3 (beta* or 1a or 1b or la or lb or "1 a" or "1 b" or "l a" or "l b"))).mp.	62
2	(multiple sclerosis or RRMS or MS).mp.	667
3	Placebo*.mp.	4772
4	And/1-3	27

Datenbank	COCHRANE Central Register of Controlled Trials	
Suchoberfläche	Ovid	
Datum der Suche	28.07.2011	
Zeitsegment	3rd Quarter 2011	
Suchfilter	--	
#	Suchbegriffe	Ergebnis
1	(extavia or betaferon or betaseron or rebif or avonex or interferonbeta* or ifnbeta* or ((peginterferon* or pegifn* or interferon* or ifn*) adj3 (beta* or 1a or 1b or la or lb or "1 a" or "1 b" or "l a" or "l b"))).mp.	1795
2	(multiple sclerosis or RRMS or MS).mp.	6192
3	placebo*.mp.	123337
4	and/1-3	214

Glatirameracetat, Abschnitt 4.3.2.1

Glatirameracetat aktiv-kontrolliert:

Datenbank	MEDLINE/MEDLINE In-Process & Other Non-Indexed Citations	
Suchoberfläche	Ovid	
Datum der Suche	21.07.2011	
Zeitsegment	1948 to July Week 1 2011 / July 20 2011	
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	interferon type I/ or Interferon Type I, Recombinant/ or Interferon-beta/ or interferons/ad, tu	22296
2	(extavia or betaferon or betaseron or rebif or avonex or interferonbeta* or ifnbeta* or ((peginterferon* or pegifn* or interferon* or ifn*) adj3 (beta* or 1a or 1b or Ia or Ib or "1 a" or "1 b" or "I a" or "I b"))).mp.	14427
3	or/1-2	28092
4	(glatiramer* or copaxon* or copolymer1 or copolymer 1 or copolymerI or copolymer I or cop 1 or cop1 or copI or cop I).mp.	1994
5	exp multiple sclerosis/	37350
6	(multiple sclerosis or RRMS or MS).ti,ab.	178345
7	or/5-6	183961
8	and/3-4,7	581
9	randomized controlled trial.pt.	311966
10	(randomized or placebo).mp.	496657
11	or/9-10	496657
12	and/8,11	105
13	(compar* or head or versus or vs).mp.	4153020
14	and/12-13	64

Datenbank	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	21.07.2011	
Zeitsegment	1980 to 2011 Week 28	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong SSL, Wilczynski NL, 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	recombinant interferon/ or recombinant beta interferon/ or beta1a interferon/ or interferon beta serine/ or beta interferon/ or beta1 interferon/	19500
2	(extavia or betaferon or betaseron or rebif or avonex or interferonbeta* or ifnbeta* or ((peginterferon* or pegifn* or interferon* or ifn*) adj3 (beta* or 1a or 1b or la or lb or "1 a" or "1 b" or "l a" or "l b"))).mp.	25120
3	or/1-2	26407
4	exp glatiramer/ or (glatiramer* or copaxon* or copolymer1 or copolymer 1 or copolymerl or copolymer l or cop 1 or cop1 or copl or cop l).mp.	4486
5	exp multiple sclerosis/	56410
6	(multiple sclerosis or RRMS or MS).ti,ab.	209607
7	or/5-6	222050
8	and/3-4,7	2391
9	(random* or double-blind*).tw.	669694
10	placebo*.mp.	251814
11	or/9-10	799265
12	And/8,11	602
13	(compar* or head or versus or vs).mp.	4175801
14	and/12-13	262

Datenbank	COCHRANE Database of Systematic Reviews	
Suchoberfläche	Ovid	
Datum der Suche	21.07.2011	
Zeitsegment	2005 to June 2011	
Suchfilter	---	
#	Suchbegriffe	Ergebnis
1	(glatiramer* or copaxon* or copolymer1 or copolymer 1 or copolymerl or copolymer l or cop 1 or cop1 or copl or cop l).mp.	13
2	(multiple sclerosis or RRMS or MS).mp.	660
3	and/1-2	12

Datenbank	COCHRANE Central Register of Controlled Trials	
Suchoberfläche	Ovid	
Datum der Suche	21.07.2011	
Zeitsegment	3rd Quarter 2011	
Suchfilter	--	
#	Suchbegriffe	Ergebnis
1	(glatiramer* or copaxon* or copolymer1 or copolymer 1 or copolymerl or copolymer I or cop 1 or cop1 or copl or cop l).mp.	152
2	(multiple sclerosis or RRMS or MS).mp.	6192
3	and/1-2	140

Glatirameracetat placebo-kontrolliert:

Datenbank	MEDLINE /MEDLINE In-Process & Other Non-Indexed Citations	
Suchoberfläche	Ovid	
Datum der Suche	20.07.2011	
Zeitsegment	1948 to July Week 1 2011 / July 19 2011	
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	(glatiramer* or copaxon* or copolymer1 or copolymer 1 or copolymerl or copolymer I or cop 1 or cop1 or copl or cop l).mp.	1993
2	exp multiple sclerosis/	37317
3	(multiple sclerosis or RRMS or MS).ti,ab.	178278
4	or/2-3	183888
5	randomized controlled trial.pt.	311606
6	(randomized or placebo).mp.	496461
7	or/5-6	496461
8	and/1,4,7	189
9	placebo*.af.	148161
10	and/8-9	103

Datenbank	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	20.07.2011	
Zeitsegment	1980 to 2011 Week 28	
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.	
1	Suchbegriffe	Ergebnis
1	exp *glatiramer/	906
2	(glatiramer* or copaxon* or copolymer1 or copolymer 1 or copolymerl or copolymer I or cop 1 or cop1 or copl or cop l).ti,ab.	2225
3	or/1-2	2439
4	exp multiple sclerosis/	56410
5	(multiple sclerosis or RRMS or MS).ti,ab.	209607
6	or/4-5	222050
7	(random* or double-blind*).tw.	669694
8	placebo*.mp.	251814
9	or/7-8	799265
10	and/3,6,9	388
11	placebo*.af.	251819
12	and/10-11	280

Datenbank	COCHRANE Database of Systematic Reviews	
Suchoberfläche	Ovid	
Datum der Suche	20.07.2011	
Zeitsegment	2005 to June 2011	
Suchfilter	---	
#	Suchbegriffe	Ergebnis
1	(glatiramer* or copaxon* or copolymer1 or copolymer 1 or copolymerl or copolymer l or cop 1 or cop1 or copl or cop l).mp.	13
2	(multiple sclerosis or RRMS or MS).mp.	660
3	and/1-2	12

Datenbank	COCHRANE Central Register of Controlled Trials	
Suchoberfläche	Ovid	
Datum der Suche	20.07.2011	
Zeitsegment	3rd Quarter 2011	
Suchfilter	--	
#	Suchbegriffe	Ergebnis
1	(glatiramer* or copaxon* or copolymer1 or copolymer 1 or copolymerl or copolymer l or cop 1 or cop1 or copl or cop l).mp.	152
2	(multiple sclerosis or RRMS or MS).mp.	6192
3	and/1-2	140

Anhang 4-B: Suchstrategien – Suche in Studienregistern

Geben Sie nachfolgend die Suchstrategien für die in Abschnitt 4.2.3.4 beschriebene Suche in Studienregistern an, und zwar einzeln für jedes Studienregister. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters (z. B. *clinicaltrials.gov*), die Internetadresse, unter der das Studienregister erreichbar ist (z. B. *http://www.clinicaltrials.gov*), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Geben Sie zusätzlich an, für welchen Abschnitt von Modul 4 die Suche durchgeführt wurde.

Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

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

Fingolimod, Abschnitt 4.3.1

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	26.07.2011
Suchstrategie	("multiple sclerosis" OR MS)[CONDITION] AND (fingolimod OR gilenya OR gilenia OR FTY720 OR "FTY 720")[INTERVENTIONS]
Treffer	20

Alle 20 Treffer sind Studien des pharmazeutischen Unternehmers.

Studienregister	clinicalstudyresults.org
Internetadresse	http://www.clinicalstudyresults.org/
Datum der Suche	26.07.2011
Suchstrategie	("Multiple Sclerosis" OR "Multiple Sclerosis, Relapsing- Remitting")[INDICATIONS] AND (("fingolimod" OR "FTY720")[GENERIC NAME] OR "Gilenya"[DRUG NAME])

Treffer	0 mit Fingolimod oder FTY720 als Prüfsubstanz
Studienregister	International Clinical Trials Registry Platform Search Portal
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	26.07.2011
Suchstrategie	(multiple sclerosis OR MS OR RRMS)[CONDITION] AND (fingolimod OR FTY720 OR "FTY 720" OR Gilenya OR Gilenia)[INTERVENTION]
Treffer	17

Alle 17 Treffer sind Studien des pharmazeutischen Unternehmers.

IFN- β -Präparate, Abschnitt 4.3.2.1

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	29.07.2011
Suchstrategie	(multiple sclerosis OR MS OR RRMS) [CONDITIONS] AND (interferon* OR betaferon OR betaseron OR extavia OR rebif OR avonex) [INTERVENTIONS]
Treffer	182

Studienregister	clinicalstudyresults.org
Internetadresse	http://www.clinicalstudyresults.org/
Datum der Suche	29.07.2011
Suchstrategie	("Multiple Sclerosis" OR "Multiple Sclerosis, Relapsing-Remitting")[INDICATIONS] AND (("Interferon Beta-1b" OR "Interferon Beta-1a")[GENERIC NAME] OR ("Betaseron" OR "Betaferon" OR "Extavia" OR "Avonex" OR "Rebif")[DRUG NAME])
Treffer	8

Studienregister	International Clinical Trials Registry Platform Search Portal
Internetadresse	http://apps.who.int/trialsearch/

Datum der Suche 29.07.2011
Suchstrategie (multiple sclerosis OR MS OR RRMS)[CONDITION] AND (Interferon% OR betaferon OR betaseron OR Avonex OR rebif OR extavia)[INTERVENTION]
Treffer 160

Glatirameracetat, Abschnitt 4.3.2.1

Studienregister clinicaltrials.gov
Internetadresse <http://www.clinicaltrials.gov>
Datum der Suche 26.07.2011
Suchstrategie (multiple sclerosis OR MS) [CONDITION] AND (glatiramer* OR copaxone OR copolymer*) [INTERVENTIONS]
Treffer 52

Studienregister clinicalstudyresults.org
Internetadresse <http://www.clinicalstudyresults.org/>
Datum der Suche 26.07.2011
Suchstrategie ("Multiple Sclerosis" OR "Multiple Sclerosis, Relapsing-Remitting")[INDICATIONS] AND (("glatiramer" OR "copolymer")[GENERIC NAME] OR "Copaxone"[DRUG NAME])
Treffer 0

Studienregister International Clinical Trials Registry Platform Search Portal
Internetadresse <http://apps.who.int/trialsearch/>
Datum der Suche 26.07.2011
Suchstrategie (multiple sclerosis OR MS OR RRMS)[CONDITION] AND (glatiramer* OR copaxone* OR copolymer*)[INTERVENTION]
Treffer 35

Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund

Listen Sie die im Volltext gesichteten und ausgeschlossenen Studien unter Nennung des Ausschlussgrunds auf.

Fingolimod-Studien (aktiv- und Placebo-kontrolliert)

Studie	Ausschlussgrund
Nicht zutreffend	

IFN- β -Studien (aktiv- und Placebo-kontrolliert)

Studie	Ausschlussgrund
Mazdeh et al., 2010	Alter der Studienpopulation ab 14 Jahre
Barbero et al., 2004	Dosis-Reduktionsstudie

Glatirameracetat-Studien (aktiv- und Placebo-kontrolliert)

Studie	Ausschlussgrund
Nicht zutreffend	

Anhang 4-D: Liste der abgebrochenen Studien

Benennen Sie in der nachfolgenden Tabelle die abgebrochenen Studien aus dem durch die verschiedenen Suchschritte in Abschnitt 4.3.1.1 identifizierten Studienpool.

Tabelle 4-85 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
Placebo-kontrolliert						
Nicht zutreffend						
aktiv-kontrolliert, zweckmäßige Vergleichstherapie(n)						
Nicht zutreffend						
a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben. b: Studie, für die der Unternehmer Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war. c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.						

Im Indikationsgebiet RRMS wurden mit dem Wirkstoff Fingolimod keine Studien abgebrochen. Eine Übersicht zu allen mit Fingolimod in der RRMS durchgeführten und laufenden Studien ist in Modul 5 hinterlegt (Modul5, Studienliste_Fingolimod.pdf).

Anhang 4-E: Liste der laufenden Studien

Benennen Sie in der nachfolgenden Tabelle die laufenden Studien aus dem durch die verschiedenen Suchschritte in Abschnitt 4.3.1.1 identifizierten Studienpool.

Tabelle 4-86 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
Placebo-kontrolliert						
D2306	nein	ja	nein	nein	ja [NCT00731692] (86;87)	nein
D2309	nein	ja	nein	nein	ja [NCT00355134] (88;89)	nein
aktiv-kontrolliert, zweckmäßige Vergleichstherapie(n)						
DUS01	nein	ja	nein	nein	ja [NCT01216072] (90;91)	nein
DIT01	nein	ja	nein	nein	ja [NCT01333501] (92;93)	nein
DIT02	nein	ja	nein	nein	ja [NCT01317004] (94;95)	nein
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben. b: Studie, für die der Unternehmer Sponsor ist oder an der der Unternehmer anderweitig finanziell beteiligt ist. c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p>						

Eine Übersicht zu allen mit Fingolimod in der RRMS durchgeführten und laufenden Studien ist in Modul 5 hinterlegt (Modul5, Studienliste_Fingolimod.pdf).

Anhang 4-F: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-87 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 4-87 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-87 (Anhang): Studiendesign und -methodik für Studie FREEDOMS

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	<p>Diese Studie war als 2-Jahresstudie entwickelt, um die Wirksamkeit, Sicherheit und Verträglichkeit zweier Dosierungen von Fingolimod (1,25 mg und 0,5 mg) im Vergleich zu Placebo bei Patienten mit RRMS zu bewerten. Sie ist Teil des klinischen Programms, das durchgeführt wird, um die Zulassung zu erhalten und Fingolimod der weltweiten klinischen Anwendung zur Verfügung zu stellen.</p> <p>Das primäre Ziel der Studie war, Überlegenheit von Fingolimod gegenüber Placebo in Hinblick auf die jährliche Schubrate zu zeigen (24 Monate). Somit war die Nullhypothese, dass kein Unterschied zwischen den beiden Interventionen bestand, zu widerlegen.</p> <p>Das wichtigste sekundäre Ziel war, die Überlegenheit von Fingolimod gegenüber Placebo hinsichtlich der Behinderungsprogression (nach 3 Monaten bestätigt) zu zeigen (24 Monate).</p> <p>Weitere sekundäre Ziele waren (24 Monate):</p> <ul style="list-style-type: none"> • Sicherheit und Verträglichkeit von Fingolimod gegenüber Placebo bei Patienten mit RRMS • Effekt von Fingolimod auf MRT Parameter der entzündlichen Krankheitsaktivität, Krankheitsausmaß und Hirnatrophie im Vergleich zu placebo • Effekt von Fingolimod auf schubbezogene Parameter im Vergleich zu Placebo (Zeit bis zum ersten Schub, Anteil schubfreier Patienten) • Effekt von Fingolimod auf die Behinderungsprogression im Vergleich zu Placebo (Zeit bis zur nach 6 Monaten bestätigten Behinderungsprogression, Anteil Patienten mit bestätigter Behinderungsprogression, Änderung des MSFC z-Scores) • Effekt von Fingolimod auf den multidimensionalen Gesundheitszustand (EQ-5D) im Vergleich zu Placebo • Pharmakokinetik von Fingolimod • PK/PD-Bezug von Fingolimod für die Hauptwirksamkeits- und Sicherheitsendpunkte <p>Explorative Ziele</p> <ul style="list-style-type: none"> • Pharmakogenetische Analysen zur Identifizierung genetischer Faktoren die mit MS in Verbindung stehen, die das Ansprechen auf Fingolimod beeinflussen, die Einfluss auf Wechselwirkungen haben oder mit einer Anfälligkeit für Nebenwirkungen in Zusammenhang stehen • Proteomisch/metabonomische Analysen • Biomarker-Analysen
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns	Dies war eine 24-monatige doppelblinde, randomisierte,

Item ^a	Charakteristikum	Studieninformation
	(z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	multizentrische, Placebo-kontrollierte Parallelgruppenstudie. Die Patienten wurden im Verhältnis 1:1:1 zu den Behandlungsarmen Fingolimod 0,5 mg, Fingolimod 1,25 mg und Placebo zugeordnet
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p>10 Amendments wurden zum Studienprotokoll erstellt. Für keines der Amendments, die nach Studienbeginn und vor Datenbankschluss effektiv wurden, wurde ein relevanter Einfluss auf die Interpretation der Studie gesehen. Folgende wesentliche Änderungen ergaben sich mit:</p> <p>Amendment 1 u.a. Auschluss diabetischer Patienten</p> <p>Nationales Amendment 1 in Polen Ausschluss von Polen für Biomarker-Analysen</p> <p>Amendment 2 u.a. Fallzahlerhöhung wegen höherer Abbrecherrate als erwartet</p> <p>Amendment 3 (nur Kanada) Ausschlusskriterium ALT/SGPT</p> <p>Amendment 4 Einführung zusätzlicher MRT-Analysen (Substudie)</p> <p>Amendment 5 u.a. zentrale Auswertung der CT Scans, zusätzliche Überwachung nach Medikationsunterbrechung wegen erhöhter Leberwerte,</p> <p>Amendment 6 u.a. Selbstuntersuchung der Haut durch die Patienten eingeführt</p> <p>Amendment 7 u.a. Maßnahmen zur Infektionsrisikominimierung eingeführt</p> <p>Amendment 8 u.a. zusätzliche Sicherheitsmaßnahmen im Nachgang zu Amendment 7 (Behördenforderung)</p> <p>Amendment 9 In Abstimmung mit EMA und FDA Anpassung der statistischen Methodik zur Auswertung des primären Endpunktes: kovarianzanalyse ersetzt durch negative Binomialregression</p>
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Generelle Einschlusskriterien:</p> <ol style="list-style-type: none"> 1. Alter ≥ 18 und ≤ 55 Jahre. 2. Frauen im gebärfähigen Alter müssen vor Eintritt in die doppelblinde Behandlungsphase einen negativen Schwangerschaftstest haben und während der Behandlung sowie 3 Monate nach deren Beendigung zwei effektive Arten der Kontrazeption anwenden (oder der Partner). Frauen, die sich bei Randomisierung 12 Monate in der Menopause befinden oder operativ sterilisiert wurden (Hysterektomie oder bilateral Oophorektomie) müssen keine Verhütungsmethode anwenden. 3. Schriftliche Zustimmung zur Studienteilnahme liegt vor. 4. MS-Diagnose entsprechend der überarbeiteten McDonald-Kriterien. 5. Schubförmig-remittierender Verlauf mit mindestens einem dokumentierten Schub im letzten oder zwei dokumentierten

Item ^a	Charakteristikum	Studieninformation
		<p>Schüben in den letzten beiden Jahren vor Randomisierung.</p> <p>6. EDSS-Wert ≥ 0 und $\leq 5,5$.</p> <p>7. Patienten, die aus welchen Gründen aus immer den Beginn oder die Fortführung Behandlung mit erhältlichen krankheitsmodifizierenden Arzneimitteln explizit ablehnen, nachdem sie von ihrem Prüfarzt über die damit verbundenen Vorzüge und mögliche unerwünschte Ereignisse aufgeklärt wurden.</p> <p>8. Neurologisch stabil ohne Hinweis auf einen Schub oder Behandlung mit Kortikosteroiden innerhalb 30 Tagen vor Randomisierung.</p> <p>Ausschlusskriterien:</p> <p>1. Andere Manifestation der MS als RRMS.</p> <p>2. Andere Autoimmunerkrankung als MS in der Vorgeschichte oder bekannter Immundefekt.</p> <p>3. Malignität aktuell oder in der Vorgeschichte (Ausnahme: Erfolgreich behandeltes Basalzellkarzinom oder Plattenepithelkarzinom der Haut).</p> <p>4. Bekannte oder neue Diagnose eines Diabetes mellitus (falls Blutzucker nüchtern ≥ 126 mg/dl oder ≥ 7 mmol/L bzw. zufällig ≥ 200 mg/dl oder $\geq 11,1$ mmol/L beim Screening wurde der Patient weitergehend auf Vorliegen eines Diabetes mellitus untersucht).</p> <p>5. Diagnose eines Makulaödems während der Prä-Randomisierungsphase (Patienten mit Makulaödem in der Vorgeschichte durften unter der Voraussetzung, dass kein Makulaödem an der ophthalmologischen Screening-Visite vorlag, teilnehmen).</p> <p>6. Aktive systemische bakterielle, virale oder fungale Infektion, Diagnose einer AIDS-, Hepatitis B oder C Infektion (definiert als positiver HIV-Antikörper-, Hepatitis-B-Oberflächenantigen- oder Hepatitis-C-Antikörpertest).</p> <p>7. Patient erhielt totallymphatische Bestrahlung oder Knochenmarktransplantation.</p> <p>8. Patient wurde mit systemischen Kortikosteroiden oder adrenokortikotropen Hormonen (ACTH) innerhalb von 1 Monat vor Randomisierung; Immunsuppressiva wie Azathioprin oder Methotrexat innerhalb von 6 Monaten vor Randomisierung; Immunglobulinen oder monoklonalen Antikörpern (einschließlich Natalizumab) innerhalb von 6 Monaten vor Randomisierung; IFN-β oder Glatirameracetat innerhalb von 3 Monaten vor Randomisierung oder irgendwann mit Cladribin, Cyclophosphamid oder Mitoxantron behandelt.</p> <p>9. Jeglicher vom primären behandelnden Arzt als medizinisch instabil beurteilter Zustand.</p> <p>10. Irgendeine der folgenden kardiovaskulären Beschwerden:</p> <ul style="list-style-type: none"> - Myokardinfarkt innerhalb von 6 Monaten vor Studieneinschluss oder aktuell instabile ischämische Herzerkrankung - Angina Pectoris infolge von Koronarspasmus in der Vorgeschichte oder Raynaud-Syndrom in der Vorgeschichte - Herzversagen zur Zeit des Screenings (Grad III nach

Item ^a	Charakteristikum	Studieninformation
		<p>NYHAC) oder sonstige schwere Herzerkrankung (vom Prüfarzt festgestellt)</p> <ul style="list-style-type: none"> - Herzstillstand, symptomatische Bradykardie, Sick-Sinus-Syndrom oder Sinoarterialer Herzblock in der Vorgeschichte oder positiver Tilt-Test zur Abklärung einer vasovagalen Synkope - Ruhepuls < 55 bpm vor Randomisierung - Aktueller oder früherer AV-Block Grad II oder Grad III oder erhöhter QTc-Intervall > 440 ms auf dem Screening-EKG - Arrhythmie, die aktuell mit Antiarrhythmika der Klasse III (z. B. Amiodaron, Bretylium, Sotalol, Ibutilid, Azimilid, Dofetilid) behandelt werden muss - Unkontrollierte Hypertonie <p>11. Irgendeine der folgenden pulmonalen Beschwerden:</p> <ul style="list-style-type: none"> - Schwere Atemwegserkrankung oder pulmonale Fibrose - Tuberkulose, außer erfolgreich behandelte Tuberkulose in der Vorgeschichte oder prophylaktische Behandlung nach positiver PPD-Hautreaktion in der Vorgeschichte - Abnorme Röntgenthorax oder hochauflösende Computertomographie (HRCT) (an ausgewählten Zentren), die auf Vorliegen einer aktiven pulmonalen Erkrankung hinweist - Abnorme Lungenfunktionstests: FEV₁, FVC-Werte < 70% der prognostizierten Werte, D_LCO-Werte < 60% der prognostizierten Werte - Asthma, das täglich (chronisch) behandelt werden muss <p>12. Irgendeine der folgenden hepatischen Beschwerden:</p> <ul style="list-style-type: none"> - Bekannter Alkoholmissbrauch, chronische Leber- oder Gallenerkrankung - Gesamtbilirubin oberhalb des Normbereichs, außer im Zusammenhang mit Gilbert-Syndrom - Konjugiertes Bilirubin oberhalb des Normbereichs - Alkalische Phosphatase mehr als 1,5 mal so hoch wie die obere Normgrenze - Aspartat-Aminotransferase (AST/SGOT), Alaninaminotransferase (ALT/SGPT) mehr als zweimal so hoch wie die obere Normgrenze (nur in Kanada: ALT/SGPT mehr als 1.5mal so hoch wie die obere Normgrenze) - Gamma-Glutamyltransferase (GGT) mehr als dreimal so hoch wie die obere Normgrenze <p>13. Einer der folgenden abnormen Laborwerte:</p> <ul style="list-style-type: none"> - Serumkreatinin > 1,7 mg/dL (150 µmol/L) - Anzahl weißer Blutkörperchen < 3,500/mm³ (< 3,5 x 10⁹/L) - Anzahl Lymphozyten < 800/mm³ (< 0,8 x 10⁹/L) <p>14. Irgendeine der folgenden neurologischen / psychischen Erkrankungen:</p> <ul style="list-style-type: none"> - Alkohol- oder Drogenmissbrauch in der Vorgeschichte oder ein anderer Faktor (z. B. schwerwiegende psychiatrische Beschwerden), die die Fähigkeit des Patienten zur Kooperation und Compliance mit dem

Item ^a	Charakteristikum	Studieninformation
		<p>Studienablauf beeinträchtigen könnte</p> <ul style="list-style-type: none"> - Progressive neurologische Erkrankung, außer MS, die die Teilnahme an der Studie beeinträchtigen könnte oder der Einnahme von Medikamenten bedarf, die vom Studienprotokoll nicht erlaubt sind <p>15. Patient ist nicht in der Lage MRT-Aufnahmen machen zu lassen, einschließlich Klaustrophobie oder Überempfindlichkeit gegenüber Gadolinium-DTPA.</p> <p>16. Teilnahme an einer anderen klinischen Forschungsstudie mit einer anderen Prüfsubstanz oder Therapie innerhalb von 6 Monaten vor Randomisierung.</p> <p>17. Schwangere oder stillende Frauen; Schwangerschaft war definiert als die Zeit zwischen Konzeption und Ende der Schwangerschaft, bestätigt mit einem positiven hCG-Labortest (> 5 mIU/mL).</p> <p>Behandlung mit Fingolimod in der Vorgeschichte.</p>
4b	Studienorganisation und Ort der Studiendurchführung	Die Studie wurde weltweit in 22 Ländern in 125 Zentren durchgeführt.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Prüfmedikation: Fingolimod wurde als Kapseln zur oralen einmal täglichen Einnahme in den Dosisstärken 1.25 mg und 0.5 mg verwendet.</p> <p>Referenztherapie: Der Prüfmedikation entsprechende Placebo-kapseln (ohne aktiven Wirkstoff).</p> <p>Die Kapseln waren ab Visite 2 (Beginn der doppelblinden Studienphase) einmal täglich, wenn möglich zur gleichen Tageszeit, mit oder ohne Mahlzeit, einzunehmen.</p>
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Das primäre Zielkriterium war die jährliche Schubrate (annualized relapse rate [ARR]) an Monat 24, was als Anzahl der Schübe pro Jahr definiert war.</p> <p>Das wichtigste sekundäre Zielkriterium war die Zeit bis zu bestätigter (nach 3 Monaten) Behinderungsprogression gemessen mit dem EDSS über 24 Monate. Bestätigte Behinderungsprogression war definiert als 1 Punkt Verschlechterung seit Baseline im EDSS-Wert falls der Baseline-Wert zwischen 0 und 5,0 lag oder eine Verschlechterung um 0,5 Punkte falls der Baseline-Wert 5,5 oder mehr betrug; die Verschlechterung musste nach 3 Monaten in Abwesenheit eines Schubes bestätigt werden.</p> <p>Weitere sekundäre Zielkriterien zur Schubrate waren der Anteil der schubfreien Patienten über 24 Monate, die Zeit bis zum ersten Schub, die Zeit bis zum zweiten Schub, die Häufigkeit der Anwendung von Kortikosteroiden zur Behandlung eines Schubes über 24 Monate, die Häufigkeit von Hospitalisierungen aufgrund eines Schubes über 24 Monate.</p> <p>Weitere sekundäre Zielkriterien zur Behinderungsprogression waren die Zeit bis zur nach 6 Monaten bestätigten Behinderungsprogression gemessen mit dem EDSS über 24 Monate, der Anteil der Patienten mit bestätigter Behinderungsprogression über 24 Monate, die Veränderung im EDSS-Wert von Baseline bis Studienende und die Veränderung im MSFC z-Punktwert von Baseline bis Studienende.</p> <p>Weitere sekundäre Zielkriterien wurden anhand MRT-Parameters gemessen:</p>

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> - Entzündungsaktivität nach 24 Monaten gemessen als die Anzahl neuer / neu vergrößerter T2-Läsionen, Anteil der Patienten ohne neue / neu vergrößerte T2-Läsionen, Anteil der Patienten ohne Gd-anreichernde T1-Läsionen, Anzahl der Gd-anreichernden T1-Läsionen, Volumen der Gd-anreichernden T1-Läsionen, Anteil der Patienten ohne neue Entzündungsaktivität (weder Gd-anreichernde T1- noch neue / neu vergrößerte T2-Läsionen) - Krankheitslast nach 24 Monaten gemessen als die Veränderung und prozentuale Veränderung des Volumens der T2-Läsionen bzw. der hypotensen T1-Läsionen ggü. Baseline • Gehirnvolumen nach 24 Monaten gemessen als prozentuale Veränderung im Gehirnvolumen ggü. Baseline (Atrophie).
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	<p>Die Zielkriterien wurden nach Studienbeginn mit Amendment 9, das am 05 Juni 2009 (vor LPLV) in Kraft trat, geändert. Nach Diskussion mit EMA und FDA wurde die Auswertung der Zielkriterien geändert und zusätzlich wurden auch die sekundären Zielkriterien vereinfacht. Als Resultat wurde als wichtigstes sekundäres Zielkriterium nur noch die nach 3 Monaten bestätigte Behinderungsprogression ausgewertet. Alle übrigen Zielkriterien wurden als sekundäre Zielkriterien analysiert.</p> <p>Das primäre Zielkriterium wurde nach Studienbeginn nicht geändert.</p>
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Die Fallzahlberechnung erfolgte für das primäre Zielkriterium (jährliche Schubrate) und das wichtigste sekundäre Zielkriterium (Zeit zur nach 3 Monaten bestätigten Behinderungsprogression an Monat 24). Die Berechnungen basieren auf dem Wilcoxon/Mann-Whitney Rangsummentest (hierarchisch um auf Multiplizität zu adjustieren).</p> <p>Die Fallzahl wurde mit Amendment 2 erhöht, um eine erhöhte Abbruchrate als angenommen zu kompensieren.</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es wurden keine Zwischenanalysen der Wirksamkeitsendpunkte durchgeführt, noch wurde die Studie vorzeitig abgebrochen.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	<p>Die Randomisierungsnummern wurden nach der folgenden Prozedur generiert wodurch sichergestellt ist, dass die Behandlungszuordnung unverzerrt und vor Patienten und ärztlichem Personal verborgen ist: Eine Randomisierungsliste wurde unter der Verantwortlichkeit des Novartis Drug Supply Management unter Verwendung eines validierten Systems hergestellt, das die zufällige Zuteilung der Behandlungsarme zu den Randomisierungsnummern im vorgegebenen Verhältnis automatisiert. Das Randomisierungsschema wurde von der Biometrics Quality Assurance Group überprüft und von ihr nach Bewilligung gesperrt.</p>
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierung wurde zentral unter Verwendung eines validierten Systems nach Studienzentrum stratifiziert durchgeführt. Pro Studienzentrum wurden die Patienten in

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

Item ^a	Charakteristikum	Studieninformation
		Blöcken randomisiert, mit einer Blockgröße von 6.
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	Die Geheimhaltung der Behandlungsfolge bis nach Datenbankschluss war gewährleistet, die Randomisierung erfolgte zentral.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Am Zentrum wurden Patienten, die die Ein-/Auschlusskriterien erfüllen, der niedrigsten, verfügbaren Nummer auf der Randomisierungsliste zugeteilt.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Alle in die Studie involvierten Personen (Patienten, Personal am Zentrum, Personal beim Sponsor außer DSM und unabhängige Mitglieder des DSMB) waren zwischen Randomisierung bis Ende der Studie verblindet. Lediglich autorisierte Personen (DSMB, DSM) hatten Zugriff zu den Behandlungs-codes.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Fingolimod und Placebo waren identisch verpackt, die Tabletten selbst waren von identischem Aussehen.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Analyse des primären Zielkriteriums: Für jeder der beiden Fingolimod-Dosen wurden die folgenden Nullhypothesen getestet: Es gibt keinen Unterschied in der jährlichen Schubrate zwischen Patienten, die mit Fingolimod behandelt wurde und Placebo-Patienten. Der Hypothesentest beruhte auf den Vergleichen Fingolimod 1,25 mg vs. Placebo und Fingolimod 0,5 mg vs. Placebo unter Verwendung des negativ binomialen Regressionsmodells. Die Kovariaten für das Modell waren Behandlung, Land, Schubrate an Baseline (in den vorigen 2 Jahren) und der Baseline-EDSS-Wert. Einzelne Länder mit kleiner Patientenzahl wurden für die Analyse gepoolt. Für die negativ binomiale Regressionsanalyse war die Responsevariable als Anzahl der Schübe pro Patient definiert und ein quadratischer Varianzschätzer wurde verwendet. Log (Zeit in der Studie in Jahren) wurde als Offsetvariable verwendet um die unterschiedliche Studiendauer der Patienten zu berücksichtigen, was das Testen der Hypothese und die Schätzung der Schubrate ermöglicht. Die jährliche Schubrate und ihr Konfidenzintervall werden von diesem Modell pro Behandlungsgruppe estimiert.</p> <p>Die primäre Analysepopulation war die ITT-Population. Als supportive Analyse wird das primäre Zielkriterium auch in der PP-Population unter Verwendung desselben Modells analysiert.</p> <p>Das primäre und das wichtigste sekundäre Zielkriterium wurden in hierarchischer Reihenfolge getestet:</p> <p>1. Fingolimod 1,25 mg vs. Placebo in Hinblick auf</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Behandlungsunterschiede für die jährliche Schubrate (negativ binomiales Modell)</p> <p>2. Fingolimod 0,5 mg vs. Placebo in Hinblick auf Behandlungsunterschiede für die jährliche Schubrate (negativ binomiales Modell)</p> <p>3. Fingolimod 1,25 mg vs. Placebo in Hinblick auf Behandlungsunterschiede für nach 3 Monaten bestätigte Behinderungsprogression (log-rank Test)</p> <p>4. Fingolimod 0,5 mg vs. Placebo in Hinblick auf Behandlungsunterschiede für nach 3 Monaten bestätigte Behinderungsprogression (log-rank Test)</p> <p>Jeder Vergleich wurde mit einem Signifikanzniveau von 0,05 getestet. Jedoch wurde ein Vergleich, der weiter unten in der Hierarchie angesiedelt war nur dann durchgeführt, wenn der jeweils höhere statistisch signifikant war. Multiplizitätsanpassung wurde durchgeführt um die Rate des Typ-1-Fehlers zu kontrollieren.</p> <p>Kategoriale MRT-Parameter wurden mit logistischer Regression analysiert, das Modell war angepasst nach Behandlung, Land und entsprechendem MRT-Baseline-Wert (falls vorhanden). Kontinuierliche MRT-Parameter und ihre Veränderung ggü. Baseline wurden hinsichtlich der Behandlungsarme mit einer ANCOVA (angepasst für Behandlung, Land und entsprechendem MRT-Baseline-Wert). Neue / neu vergrößerte T2-Läsionen wurden zwischen den Behandlungsarmen mit dem negativ binomialen Modell, angepasst für Behandlung und Land, verglichen. Für die Analyse der Zeit bis zum ersten bzw. zweiten Schub wurden die Überlebenskurven mit dem Log-Rank-Test zwischen Fingolimod und Placebo verglichen. Als supportive Analyse wurde das Cox proportional hazards Modell verwendet um die Zeit bis zum Event zu modellieren, adjustiert für Behandlung, Land, Anzahl der Schübe in der vorangegangenen 2 Jahren und Baseline-EDSS-Wert.</p> <p>Zusätzlich wurden die Kaplan-Meier-Schätzer an 12 und 24 Monaten inklusive der 95%-Konfidenzintervalle präsentiert und die entsprechenden Kaplan-Meier-Kurven pro Behandlungsgruppe gebildet.</p> <p>Time-to-Event-Analysen zu weiteren Variablen der Behinderungsprogression wurden wie das wichtigste sekundäre Zielkriterium ausgewertet.</p> <p>Veränderung im EDSS-Wert und MSFC z-Punktwert ggü. Baseline wurden mittels ANCOVA (adjustiert für Behandlung, Land, entsprechender Baseline-Wert und Alter) um die Werte zwischen den Behandlungsarmen zu vergleichen.</p> <p>Weitere kategoriale und kontinuierliche sekundäre Variablen wurden zusammengefasst und mit dem Fisher Exact-Test bzw. dem Wilcoxon-Rangsummentest verglichen.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Keine weiteren Analysen waren im Studienprotokoll definiert. Post-hoc wurden jedoch die folgenden Subgruppenanalysen mit dem primären und dem wichtigsten sekundären Zielkriterium durchgeführt: Geschlecht (weiblich, männlich), Alter (\leq Median, $>$ Median in Jahren), Anzahl der Schübe in den letzten zwei Jahren (0, 1, 2-3, 4-5, $>$ 5 Schübe), MS Vorgeschichte (ein Schub, mehrere Schübe), Baseline-EDSS-Wert ($<$ 3,0, \geq 3,0), Dauer der MS seit dem ersten Symptom (\leq Median, $>$ Median in Jahren), frühere MS krankheitsmodifizierende Arzneimittel

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

Item ^a	Charakteristikum	Studieninformation
		(keine Vorbehandlung, Vorbehandlung) und Anzahl der Gd-anreichernden T1-Läsionen an Baseline (0, 1-2, ≥ 3).
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Fingolimod 1,25 mg a) randomisiert wurden: n= 429 b) davon behandelt: n= 429 c) primäre Analyse: n= 429 (ITT) Fingolimod 0,5 mg a) randomisiert wurden: n= 425 b) davon behandelt: n= 425 c) primäre Analyse: n= 425 (ITT) Placebo a) randomisiert wurden: n=418 b) davon behandelt: n=418 c) primäre Analyse: n=418 (ITT)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Fingolimod 1,25 mg: Folgende Gründe führten zum Medikationsabbruch nach Randomisierung Widerruf der Einwilligungserklärung: n=30 Unerwünschtes Ereignis: n=31 Unbefriedigender therap. Effekt: n=18 Abnormer Laborwert: n=32 Erschien nicht mehr zu den Visiten: n=2 Administrative Probleme: n=3 Protokollverletzung: n=8 Abnormes Testergebnis: n=6 Todesfall: n=1 Fingolimod 0,5 mg: Folgende Gründe führten zum Medikationsabbruch nach Randomisierung Widerruf der Einwilligungserklärung: n=17 Unerwünschtes Ereignis: n=15 Unbefriedigender therap. Effekt: n=8 Abnormer Laborwert: n=20 Erschien nicht mehr zu den Visiten: n=6 Administrative Probleme: n=3 Protokollverletzung: n=8 Abnormes Testergebnis: n=3 Placebo: Folgende Gründe führten zum Medikationsabbruch nach Randomisierung Widerruf der Einwilligungserklärung: n=31 Unerwünschtes Ereignis: n=24 Unbefriedigender therap. Effekt: n=36 Abnormer Laborwert: n=5 Erschien nicht mehr zu den Visiten: n=5 Administrative Probleme: n=4 Protokollverletzung: n=5 Abnormes Testergebnis: n=3 Todesfall: n=2
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der	FPFV: 26. Januar 2006 LPLV: 29. Juli 2009

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

Item ^a	Charakteristikum	Studieninformation
	Nachbeobachtung	
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie endet planmäßig nach der letzten Visite des letzten in der Studie behandelten Patienten. Das Rekrutierungsziel war erreicht.
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

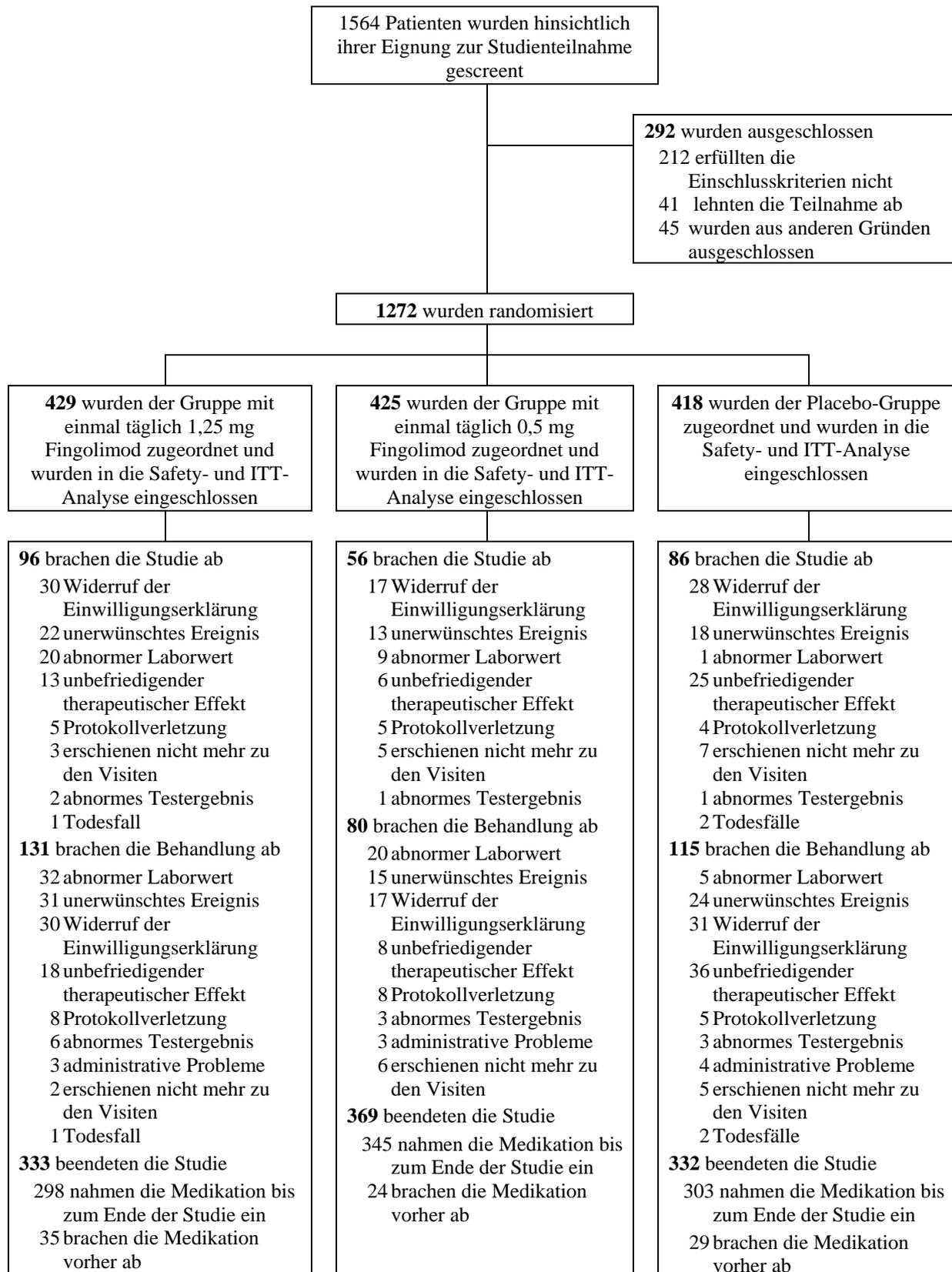


Tabelle 4-88 (Anhang): Studiendesign und -methodik für Studie TRANSFORMS

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	<p>Diese Studie wurde entwickelt um Daten zur Wirksamkeit, Sicherheit und Verträglichkeit von Fingolimod im Vergleich zu IFN-β-1a (Avonex®) bei Patienten mit RRMS zu erbringen sowie Langzeitdaten zu Sicherheit und Wirksamkeit von Fingolimod während der optional Extensionsphase der Studie. Die Studie ist Teil des klinischen Programms, das durchgeführt wird, um die Zulassung zu erhalten und Fingolimod der weltweiten klinischen Anwendung zur Verfügung zu stellen. Die Studie wurde global durchgeführt.</p> <p>Das primäre Ziel der Studie war Überlegenheit von Fingolimod gegenüber IFN-β-1a in Hinblick auf die jährliche Schubrate zu bei Patienten mit RRMS, die über 12 Monate behandelt wurden zu zeigen. Somit war die Nullhypothese, dass kein Unterschied zwischen den beiden Interventionen bestand, zu widerlegen.</p> <p>Das wichtigste sekundäre Ziel war, Überlegenheit von Fingolimod gegenüber IFN-β-1a i. m. hinsichtlich des Effekts auf die entzündliche Krankheitsaktivität (T2-Läsionen) und die Behinderungsprogression (nach 3 Monaten bestätigt)</p> <p>Weitere sekundäre Ziele waren (12 Monate):</p> <ul style="list-style-type: none"> • Sicherheit und Verträglichkeit von Fingolimod gegenüber IFN-β-1a i. m. bei Patienten mit RRMS • Effekt von Fingolimod auf schubbezogene Parameter im Vergleich zu IFN-β-1a i. m. (Zeit bis zum ersten Schub, Anteil schubfreier Patienten) • Effekt von Fingolimod auf MRT Parameter der entzündlichen Krankheitsaktivität, Krankheitsausmaß und Gewebeerstörung im Vergleich zu IFN-β-1a i. m. • PK/PD-Bezug von Fingolimod für die Hauptwirksamkeits- und Sicherheitsendpunkte <p>Weitere Ziele</p> <ul style="list-style-type: none"> • Die Effekte von Fingolimod auf Patientenberichtete Endpunkte (gesundheitsbezogene Lebensqualität, Aktivitäten des täglichen Lebens, Fatigue) • Messung der Hirnatrophie • Pharmakogenetische Analysen zur Identifizierung genetischer Faktoren die mit MS in Verbindung stehen, die das Ansprechen auf Fingolimod beeinflussen, die Einfluss auf Wechselwirkungen haben oder mit einer Anfälligkeit für Nebenwirkungen in Zusammenhang stehen • Proteomisch/metabonomische Analysen
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Dies war eine doppelblinde, randomisierte, multizentrische, aktiv kontrollierte Parallelgruppenstudie über 12 Monate. Die Patienten wurden im Verhältnis 1:1:1 zu den Behandlungsarmen Fingolimod 0,5 mg, Fingolimod 1,25 mg und IFN- β -1a

Item ^a	Charakteristikum	Studieninformation
		zugeordnet.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p>10 Amendments modifizierten das Protokoll, für keines wurde ein relevanter Einfluss auf die Interpretation der Studie gesehen. Die wesentlichen Inhalte waren:</p> <p>Amendment 1 Klärungen zur Extensionsphase</p> <p>Amendment 2 (nur Kanada) u.a. Anpassung der Ausschlusskriterien hinsichtlich der Leberwerte, zusätzliche Überwachung bei Patienten, die aufgrund von Leberwerterhöhungen die Therapie unterbrochen hatten</p> <p>Amendment 3 u.a. Klarstellungen, Einführung zusätzlicher Sicherheitsmaßnahmen</p> <p>Amendment 4 (nur USA) Spezifizierungen zu ophthalmologischen Untersuchungen und Abbruchkriterien</p> <p>Amendment 5 (nur Ungarn) Abkopplung der Extensionsphase</p> <p>Amendment 6 u.a. Anpassung des Hintergrundkapitels und der Patienteninformation bezüglich aufgetretener Makulaödem-Fälle, Einführung strengerer Kriterien zur Schwangerschaftsverhütung, Spezifizierung und Einführung von Sicherheitsmaßnahmen, v.a. bezüglich der Krebserkennug</p> <p>Amendment 7 Einführung einer Selbstuntersuchung der Haut durch die Patienten</p> <p>Amendment 8 u.a. erneute Anleitung zum Umgang mit und Prävention von Infektionen, Anleitung zum Umgang mit Makulaödem</p> <p>Amendment 9 In Absprache mit EMA und FDA wurde für den primären Endpunkt und den wichtigsten sekundären Endpunkt eine Adjutierung auf Multiplizität eingeführt.</p> <p>Amendment 10 u.a. zusätzliche Sicherheitsmaßnahmen im Nachgang zu Amendment 8, Änderung der Grenzwerte für Lymphopenien</p>
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Generelle Einschlusskriterien:</p> <p>1. Männliche oder weibliche Teilnehmer. Frauen im gebärfähigen Alter müssen</p> <ul style="list-style-type: none"> - vor Eintritt in die doppelblinde Behandlungsphase einen negativen Schwangerschaftstest haben - während der Behandlung sowie 3 Monate nach deren Beendigung zwei effektive Arten der Kontrazeption anwenden (oder der Partner). <p>Frauen, die sich bei Randomisierung 12 Monate in der Menopause befinden oder operativ sterilisiert wurden (Hysterektomie oder bilateral Oophorektomie) müssen keine Verhütungsmethode anwenden.</p>

Item ^a	Charakteristikum	Studieninformation
		<p>2. Alter ≥ 18 und ≤ 55 Jahre.</p> <p>3. Schriftliche Zustimmung zur Studienteilnahme liegt vor.</p> <p>Multiple Sklerose:</p> <p>4. MS-Diagnose entsprechend der 2005 überarbeiteten McDonald-Kriterien.</p> <p>5. Schubförmig-remittierender Verlauf mit mindestens einem dokumentierten Schub im letzten oder zwei dokumentierten Schüben in den letzten beiden Jahren vor Randomisierung.</p> <p>6. EDSS-Wert ≥ 0 und $\leq 5,5$.</p> <p>7. Neurologisch stabil ohne Hinweis auf einen Schub oder Behandlung mit Kortikosteroiden innerhalb 30 Tagen vor Randomisierung.</p> <p>Ausschlusskriterien:</p> <p>1. Andere Manifestation der MS als RRMS.</p> <p>2. Andere Autoimmunerkrankung als MS in der Vorgeschichte oder bekannter Immundefekt.</p> <p>3. Epileptische Anfälle innerhalb von 3 Monaten vor Randomisierung</p> <p>4. Malignität aktuell oder in der Vorgeschichte (Ausnahme: Erfolgreich behandeltes Basalzellkarzinom oder Plattenepithelkarzinom der Haut).</p> <p>5. Bekannte oder neue Diagnose eines Diabetes mellitus (falls Blutzucker nüchtern ≥ 126 mg/dl oder ≥ 7 mmol/L bzw. zufällig ≥ 200 mg/dl oder $\geq 11,1$ mmol/L beim Screening wurde der Patient weitergehend auf Vorliegen eines Diabetes mellitus untersucht).</p> <p>6. Diagnose eines Makulaödems während der Prä-Randomisierungsphase (Patienten mit Makulaödem in der Vorgeschichte durften unter der Voraussetzung, dass kein Makulaödem an der ophthalmologischen Screening-Visite vorlag, teilnehmen).</p> <p>7. Aktive systemische bakterielle, virale oder fungale Infektion, Diagnose einer AIDS-, Hepatitis B oder C Infektion (definiert als positiver HIV-Antikörper-, Hepatitis-B-Oberflächenantigen- oder Hepatitis-C-Antikörpertest).</p> <p>8. Patient erhielt totallymphatische Bestrahlung oder Knochenmarktransplantation.</p> <p>9. Patient wurde mit systemischen Kortikosteroiden oder adrenokortikotropen Hormonen (ACTH) innerhalb von 1 Monat vor Randomisierung; Immunsuppressiva wie Azathioprin oder Methotrexat innerhalb von 6 Monaten vor Randomisierung; Immunglobulinen oder monoklonalen Antikörpern (einschließlich Natalizumab) innerhalb von 6 Monaten vor Randomisierung oder irgendwann mit Cladribin, Cyclophosphamid oder Mitoxantron behandelt.</p> <p>10. Jeglicher vom primären behandelnden Arzt als medizinisch instabil beurteilter Zustand.</p> <p>11. Irgendeine der folgenden kardiovaskulären Beschwerden:</p> <ul style="list-style-type: none"> - Myokardinfarkt innerhalb von 6 Monaten vor Studieneinschluss oder aktuell instabile ischämische Herzerkrankung - Angina Pectoris infolge von Koronarspasmus in der Vorgeschichte oder Raynaud-Syndrom in der

Item ^a	Charakteristikum	Studieninformation
		<p>Vorgeschichte</p> <ul style="list-style-type: none"> - Herzversagen zur Zeit des Screenings (Grad III nach NYHAC) oder sonstige schwere Herzerkrankung (vom Prüfarzt festgestellt) - Herzstillstand in der Vorgeschichte, symptomatische Bradykardie, Sick-Sinus-Syndrom oder Sinoarterialer Herzblock in der Vorgeschichte oder positiver Tilt-Test zur Abklärung einer vasovagalen Synkope - Ruhepuls < 55 bpm vor Randomisierung - Aktueller oder früherer AV-Block Grad II oder Grad III oder erhöhter QTc-Intervall > 440 ms auf dem Screening-EKG - Arrhythmie, die aktuell mit Antiarrhythmika der Klasse III (z. B. Amiodaron, Bretylium, Sotalol, Ibutilid, Azimilid, Dofetilid) behandelt werden muss - Unkontrollierte Hypertonie <p>12. Irgendeine der folgenden pulmonalen Beschwerden:</p> <ul style="list-style-type: none"> - Schwere Atemwegserkrankung oder pulmonale Fibrose - Tuberkulose, außer erfolgreich behandelte Tuberkulose in der Vorgeschichte oder prophylaktische Behandlung nach positiver PPD-Hautreaktion in der Vorgeschichte - Abnorme hochauflösende Computertomographie (HRCT) (oder Röntgenthorax falls HRCT aufgrund lokaler Verordnungen nicht erlaubt ist), die auf Vorliegen einer aktiven pulmonalen Erkrankung hinweist - Abnorme Lungenfunktionstests: FEV₁, FVC-Werte < 70% der prognostizierten Werte, D_LCO-Werte < 60% der prognostizierten Werte - Chronische Asthmatherapie <p>13. Irgendeine der folgenden hepatischen Beschwerden:</p> <ul style="list-style-type: none"> - Bekannter Alkoholmissbrauch, chronische Leber- oder Gallenerkrankung - Gesamtbilirubin oberhalb des Normbereichs, außer im Zusammenhang mit Gilbert-Syndrom - Konjugiertes Bilirubin oberhalb des Normbereichs - Alkalische Phosphatase mehr als 1,5 mal so hoch wie die obere Normgrenze - Aspartat-Aminotransferase (AST), Alaninaminotransferase (ALT) mehr als zweimal so hoch wie die obere Normgrenze - Gamma-Glutamyltransferase (GGT) mehr als dreimal so hoch wie die obere Normgrenze <p>14. Einer der folgenden abnormen Laborwerte:</p> <ul style="list-style-type: none"> - Serumkreatinin > 1,7 mg/dL (150 µmol/L) - Anzahl weißer Blutkörperchen < 3,500/mm³ (< 3,5 x 10⁹/L) - Anzahl Lymphozyten < 800/mm³ (< 0,8 x 10⁹/L) <p>15. Irgendeine der folgenden neurologischen / psychischen Erkrankungen:</p> <ul style="list-style-type: none"> - Schwere depressive Episode innerhalb von 3 Monaten vor Randomisierung - Relevante Vorgeschichte eines Selbstmordversuchs oder

Item ^a	Charakteristikum	Studieninformation
		<p>der Patient ist selbstmordgefährdet nach Ansicht des Prüfarztes</p> <ul style="list-style-type: none"> - Alkohol- oder Drogenmissbrauch in der Vorgeschichte oder ein anderer Faktor (z. B. schwerwiegende psychiatrische Beschwerden), die die Fähigkeit des Patienten zur Kooperation und Compliance mit dem Studienablauf beeinträchtigen könnte - Progressive neurologische Erkrankung, außer MS, die die Teilnahme an der Studie beeinträchtigen könnte oder der Einnahme von Medikamenten bedarf, die vom Studienprotokoll nicht erlaubt sind <p>16. Patient ist nicht in der Lage MRT-Aufnahmen machen zu lassen, einschließlich Klaustrophobie oder Überempfindlichkeit gegenüber Gadolinium-DTPA.</p> <p>17. Überempfindlichkeit ggü. natürlichem oder rekombinantem Interferon beta, Humanalbumin oder einem anderen Bestandteil der Rezeptur.</p> <p>18. Teilnahme an einer anderen klinischen Forschungsstudie mit einer anderen Prüfsubstanz oder Therapie innerhalb von 6 Monaten vor Randomisierung.</p> <p>Behandlung mit Fingolimod in der Vorgeschichte.</p>
4b	Studienorganisation und Ort der Studiendurchführung	Die Studie wurde weltweit in 18 Ländern durchgeführt.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Prüfmedikation: Fingolimod wurde als Kapseln zur oralen einmal täglichen Einnahme in den Dosisstärken 1.25 mg und 0.5 mg verwendet.</p> <p>Aktive Kontrolle: INFβ-1a 30µg in Fertigspritzen zur intramuskulären (i. m.) Injektion einmal wöchentlich</p> <p>Referenztherapie: Der Prüfmedikation entsprechende Placebokapseln (ohne aktiven Wirkstoff) sowie der aktiven Kontrolle entsprechende Placebofertigspritzen (ohne aktiven Wirkstoff).</p> <p>Die Kapseln (Fingolimod und Placebo) waren ab Visite 3 (Beginn der doppelblinden Studienphase) einmal täglich, wenn möglich zur gleichen Tageszeit, mit oder ohne Mahlzeit, einzunehmen. Die Spritzen (INFβ-1a und Placebo) waren ab Visite 3 einmal wöchentlich intermuskulär zu verabreichen.</p> <p>Alle Patienten erhielten sowohl Tabletten als auch Fertigspritzen (double-dummy Design). Die Patienten wurden vom Zentrumspersonal hinsichtlich der Verabreichung der Spritzen instruiert und trainiert.</p>
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Das primäre Zielkriterium war die jährliche Schubrate (annualized relapse rate [ARR]), definiert als Anzahl der Schübe pro Jahr.</p> <p>Die zwei wichtigsten sekundären Kriterien waren:</p> <ul style="list-style-type: none"> - Die Wirkung auf die entzündliche Erkrankung gemessen anhand der neuen / neu vergrößerten T2-Läsionen am Monat 12. - Die Zeit bis zu bestätigter (nach 3 Monaten) Behinderungsprogression gemessen mit dem EDSS über 12 Monate. Bestätigte Behinderungsprogression war definiert als 1 Punkt Verschlechterung seit Baseline im EDSS-Wert falls der Baseline-Wert zwischen 0 und 5,0 lag oder eine Verschlechterung um 0,5 Punkte falls der Baseline-Wert 5,5

Item ^a	Charakteristikum	Studieninformation
		<p>oder mehr betrug; die Verschlechterung musste nach 3 Monaten in Abwesenheit eines Schubes bestätigt werden.</p> <ul style="list-style-type: none"> • Weitere sekundäre Zielkriterien waren die Anzahl der Schübe (bestätigte und unbestätigte), der Anteil der schubfreien Patienten, Anteil der Patienten mit nach 3 Monaten bestätigter Behinderungsprogression, die Zeit bis zu bestätigter Behinderungsprogression, die seit der letzten Beobachtung besteht, die Zeit zu schwerer Behinderung (EDSS \geq 6,0), der EDSS-Wert, der MSFC-z-Punktwert, die MSFC-Subskalen, der Anteil der Patienten ohne neue / neu vergrößerte T2-Läsionen an Monat 12, Anteil der Patienten ohne Gd-anreichernde T1-Läsionen, Anzahl der Gd-anreichernden T1-Läsionen, Volumen der Gd-anreichernden T1-Läsionen, Anteil der Patienten ohne neue Entzündungsaktivität (weder Gd-anreichernde T1- noch neue / neu vergrößerte T2-Läsionen, Veränderung und prozentuale Veränderung des Volumens der T2-Läsionen bzw. der hypotensen T1-Läsionen ggü. Baseline an Monat 12, das normierte Gehirnvolumen an Baseline und prozentuale Veränderung im Gehirnvolumen ggü. Baseline an Monat 12 (gemessen anhand struktureller Bildevaluierung unter Verwendung der SIENA-Methode [normalization of atrophy]).
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	<p>Die Zielkriterien wurden nach Studienbeginn mit Amendment 9, das am 25 Juli 2008 (vor LPLV) in Kraft trat, geändert. Nach Diskussion mit EMA und FDA wurde die Auswertung der Zielkriterien geändert. Als wichtigste sekundäre Zielkriterien wurden nach Diskussion mit FDA nur noch die Anzahl neuer / neu vergrößerter T2-Läsionen und die nach 3 Monaten bestätigte Behinderungsprogression ausgewertet. Alle übrigen Zielkriterien wurden als sekundäre Zielkriterien analysiert. Das primäre Zielkriterium wurde nach Studienbeginn nicht geändert.</p>
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Zur Fallzahlberechnung wurde der Wilcoxon / Mann-Whitney-Rangsummentest zum Vergleich der Fingolimod 1,25 mg-Gruppe mit der INFβ-1a-Gruppe. In der Studie CFTY720D2201 wurde eine relative Reduktion der jährlichen Schubrate von 54,5% in der Fingolimod 1,25 mg-Gruppe im Vergleich zur Placebo-Gruppe gefunden. Basierend auf historischen Daten zu INFβ-1a i. m. und möglicher Unterschiede zwischen den Patientenpopulationen wurden jährliche Schubraten von 0,55 und 0,33 für INFβ-1a i. m. bzw. Fingolimod 1,25 mg angenommen (relative Reduktion 40%). Basierend auf Daten aus der Studie CFTY720D2201, ihrer Extension und begrenzten historischen Daten mit anderen MS-Therapien wurde eine allgemeine Standardabweichung von 0,9 angenommen. Unter diesen Annahmen würden 368 Patienten pro Gruppe eine 90%ige Power bei einem zweiseitigen Signifikanzniveau von 0,05 erbringen.</p> <p>In der Studie CFTY720D2201 betrug die Abbrecherrate pro halbem Jahr ca. 8%. Extrapoliert man diesen Wert auf die 12-monatige Studie und nimmt an, dass diese Patienten nicht in den Behandlungsvergleich einbezogen werden können, werden 57 Patienten (15,5%) pro Gruppe hinzugefügt. Somit ergibt sich eine Gesamtfallzahl von 1275 benötigten Patienten (425 Patienten pro Gruppe).</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Basierend auf einer Fallzahl von 425 Patienten pro Gruppe wurde die Power für die Analyse der beiden wichtigsten sekundären Kriterien berechnet:</p> <ol style="list-style-type: none"> 1. Mit der errechneten Stichprobenzahl von 368 Patienten, die die Studie beenden, ergibt sich eine 90%ige Power für die Ermittlung eines Therapieunterschieds zwischen Fingolimod 1,25 mg vs. INFβ-1a i. m. unter Verwendung eines konservativen Wilcoxon Rangsummentests mit einem zweiseitigen Signifikanzniveau von 0,05. 2. Mit 425 randomisierten Patienten und 57 Abbrechern pro Gruppe (exponentiell verteilt) und unter der Annahme einer 12-Monats-Progressionsrate für die Fingolimod-Gruppe von 12% (eine relative Reduktion von 20% von INFβ-1a i. m.) ergibt sich eine 22%ige Power zur Ermittlung eines Therapieunterschieds unter Verwendung des Log-Rank-Test mit einem zweiseitigen Signifikanzniveau von 0,05. <p>Die Fallzahlberechnung wurde anhand der kommerziell erhältlichen Software nQuery (Version 5.0) durchgeführt.</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es wurden keine Zwischenanalysen der Wirksamkeitsendpunkte durchgeführt, noch wurde die Studie vorzeitig abgebrochen.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Eine Randomisierungsliste wurde vom IVRS-Anbieter unter Verwendung eines validierten Systems hergestellt. Die Liste ordnet jedem Patienten eine Randomisierungsnummer zu, die mit einem Behandlungsarm und wiederum mit einer Medikationsnummer verlinkt ist. Eine separate Medikationsrandomisierungsliste wurde unter der Verantwortlichkeit des Novartis Drug Supply Management unter Verwendung eines validierten Systems hergestellt, das die zufällige Zuteilung der Behandlungsarme zu den Randomisierungsnummern im vorgegebenen Verhältnis automatisiert. Das Randomisierungsschema wurde von der Biometrics Quality Assurance Group überprüft und von ihr nach Bewilligung gesperrt.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierung wurde zentral unter Verwendung eines validierten Systems nach Studienzentrum stratifiziert durchgeführt. Pro Studienzentrum wurden die Patienten in Blöcken randomisiert, mit einer Blockgröße von 6.
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	Die Randomisierung wurde zentral über ein Interaktives Voice-Response-System (IVRS) durchgeführt. Die Randomisierungs- und Medikationslisten waren so erstellt worden, dass durch den Anruf beim IVRS der Patient zwar eindeutig einer Randomisierungs- und Medikationsnummer und somit auch einer Therapie zugeordnet war, dies jedoch dem Anrufer verborgen blieb. Die Geheimhaltung der Randomcodes wurde bis Datenbankschluss aufrechterhalten.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die	Die Randomisierung wurde zentral über ein Interaktives Voice-Response-System (IVRS) durchgeführt. Dazu rief das entsprechende Zentrum wenn ein Patient für die Studie randomisiert werden sollte, das IVRS an, bestätigte, dass die Ein- und Ausschlusskriterien erfüllt waren und erhielt darüber

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

Item ^a	Charakteristikum	Studieninformation
	Probanden/Patienten den Gruppen zu?	die Randomisierungsnummer für diesen Patienten. Die Randomisierungsnummer ordnete den Patienten einem Behandlungsarm sowie einer eindeutig nummerierten Packung an Studienmedikation zu, ohne dass der Anrufer die Nummer der Packung erfährt. Diese ist über die Randomisierungsliste eindeutig einer Randomisierungsnummer zugeordnet.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Alle in die Studie involvierten Personen (Patienten, Personal am Zentrum, Personal beim Sponsor außer DSM und unabhängige Mitglieder des DSMB) waren zwischen Randomisierung bis Ende der Studie verblindet. Lediglich autorisierte Personen (DSMB, DSM) hatten Zugriff zu den Behandlungs-codes.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Fingolimod und entsprechendes Placebo waren identisch verpackt, die Tabletten waren von identischem Aussehen. INFβ-1a und entsprechende Placebo-Spritzen waren ebenfalls identisch verpackt, die Fertigspritzen waren von identischem Aussehen.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Analyse des primären Zielkriteriums: Für jeder der beiden Fingolimod-Dosen wurden die folgenden Nullhypothesen getestet: Es gibt keinen Unterschied in der jährlichen Schubrate zwischen Patienten, die mit Fingolimod behandelt wurde und INFβ-1a i. m.-Patienten. Der Hypothesentest beruhte auf den Vergleichen Fingolimod 1,25 mg vs. INFβ-1a i. m. und Fingolimod 0,5 mg vs. INFβ-1a i. m. unter Verwendung des negativ binomialen Regressionsmodells. Die Kovariaten für das Modell waren Behandlung, Land, Schubrate an Baseline (in den vorigen 2 Jahren) und der Baseline-EDSS-Wert.</p> <p>Für die negativ binomiale Regressionsanalyse war die Responsevariable als Anzahl der Schübe pro Patient definiert und ein quadratischer Varianzschätzer wurde verwendet. Log (Zeit in der Studie in Jahren) wurde als Offsetvariable verwendet um die unterschiedliche Studiendauer der Patienten zu berücksichtigen, was das Testen der Hypothese und die Schätzung der Schubrate ermöglicht. Die jährliche Schubrate und ihr Konfidenzintervall werden von diesem Modell pro Behandlungsgruppe estimiert. Das Verhältnis der jährlichen Schubrate zwischen den Gruppen und das dazugehörige 95%-KI wurde ebenfalls ausgegeben.</p> <p>Als supportive Analyse wird das primäre Zielkriterium auch in der PP-Population unter Verwendung desselben Modells analysiert. Zusätzlich wurde als supportive Analyse eine ANCOVA mit Behandlungsgruppe, Land, Anzahl der Schübe in den letzten 2 Jahren (Baseline-Wert) und Baseline-EDSS-Wert als Kovariaten gerechnet.</p> <p>Das primäre und die wichtigsten sekundären Zielkriterien wurden in hierarchischer Reihenfolge getestet:</p> <ol style="list-style-type: none"> 1. Fingolimod 1,25 mg vs. INFβ-1a i. m. in Hinblick auf Behandlungsunterschiede für die jährliche Schubrate (negativ binomiales Modell) 2. Fingolimod 0,5 mg vs. INFβ-1a i. m. in Hinblick auf Behandlungsunterschiede für neue /neu vergrößerte T2-

Item ^a	Charakteristikum	Studieninformation
		<p>Läsionen (negativ binomiales Modell)</p> <p>3. Fingolimod 1,25 mg vs. INF-β-1a i. m. in Hinblick auf Behandlungsunterschiede für neue /neu vergrößerte T2-Läsionen (negativ binomiales Modell)</p> <p>4. Fingolimod 0,5 mg vs. INF-β-1a i. m. in Hinblick auf Behandlungsunterschiede für die jährliche Schubrate (negativ binomiales Modell)</p> <p>5. Fingolimod 1,25 mg vs. INF-β-1a i. m. in Hinblick auf Behandlungsunterschiede für nach 3 Monaten bestätigte Behinderungsprogression (cox proportional hazard Modell, log-rank Test)</p> <p>6. Fingolimod 0,5 mg vs. INF-β-1a i. m. in Hinblick auf Behandlungsunterschiede für nach 3 Monaten bestätigte Behinderungsprogression (cox proportional hazard Modell, log-rank Test)</p> <p>Jeder Vergleich wurde mit einem Signifikanzniveau von 0,05 getestet. Jedoch wurde ein Vergleich, der weiter unten in der Hierarchie angesiedelt war nur dann durchgeführt, wenn der jeweils höhere statistisch signifikant war. Multiplizitätsanpassung wurde durchgeführt um die Rate des Typ-1-Fehlers zu kontrollieren.</p> <p>Kategoriale MRT-Parameter wurden mit logistischer Regression analysiert, das Modell war angepasst nach Behandlung, Land und entsprechendem MRT-Baseline-Wert (falls vorhanden). Kontinuierliche MRT-Parameter und ihre Veränderung ggü. Baseline wurden hinsichtlich der Behandlungsarme mit einer ANCOVA (angepasst für Behandlung, Land und entsprechendem MRT-Baseline-Wert). Für die Analyse der Zeit bis zum ersten bzw. zweiten Schub wurden mit Cox Regressionsanalyse mit den gleichen Kovariaten wie sie in der primären Analyse verwendet wurden zwischen Fingolimod und INFβ-1a i. m. verglichen. Der Anteil der schubfreien Patienten wurde mit logistischer Regression (adjustiert für Land, Baseline-Schubrate in den letzten 2 Jahren und Baseline-EDSS-Wert) zwischen den Behandlungsgruppen verglichen. Sowohl die naiven als auch die Kaplan-Meier-Schätzer für time-to-event-Analysen an 12 Monaten wurden angegeben. Die beiden Fingolimod-Gruppen wurden hinsichtlich der jährlichen Schubrate mit dem negativ binomialen Modell das in der primären Analyse verwendet wurde, verglichen.</p> <p>Kategoriale und kontinuierliche Variablen (Veränderungen im EDSS-Wert und MSFC z-Punktwert) zur Behinderungsprogression wurden mit dem Wilcoxon-Rangsummentest bzw. einer ANOCOVA adjustiert für Behandlung, Land und entsprechendem MRT-Wert (falls vorhanden). Proportionen wurden mit einem logistischen Regressionsmodell adjustiert für Behandlung, Land und entsprechendem MRT-Wert analysiert Sowohl die naiven als auch die Kaplan-Meier-Schätzer für time-to-event-Analysen an 12 Monaten wurden angegeben.</p> <p>Weitere sekundäre Variablen zur Lebensqualität wurden deskriptiv beschrieben bzw. mit dem Wilcoxon-Rangsummentest und dem bereits beschriebenen ANOCOVA-Modell im Hinblick auf Gruppenunterschiede analysiert.</p>

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

12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Keine weiteren Analysen waren im Studienprotokoll definiert. Post hoc Subgruppenanalysen wurden zum primären Zielkriterium hinsichtlich Geschlecht (weiblich, männlich), Alter (\leq Median, $<$ Median in Jahren), Vorbehandlung (ja, nein) und EDSS-Baseline-Wert (0,0-2,5, 3,0-5,0, \geq 5,5) durchgeführt.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Fingolimod 1,25 mg a) randomisiert wurden: n= 426 b) davon behandelt: n= 420 c) primäre Analyse: n= 420 (ITT) Fingolimod 0,5 mg a) randomisiert wurden: n= 431 b) davon behandelt: n= 429 c) primäre Analyse: n= 429 (ITT) INF β -1a i. m. a) randomisiert wurden: n=435 b) davon behandelt: n=431 c) primäre Analyse: n=431 (ITT)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Fingolimod 1,25 mg: Folgende Gründe führten zum Medikationsabbruch nach Randomisierung Widerruf der Einwilligungserklärung: n=10 Unerwünschtes Ereignis: n=32 Unbefriedigender therap. Effekt: n=5 Abnormer Laborwert: n=8 Erschien nicht mehr zu den Visiten: n=1 Administrative Probleme: n=1 Abnormes Testergebnis: n=3 Zustand bedurfte keiner weiteren Behandlung mit Studienmedikation: n=1 Todesfall: n=1 Fingolimod 0,5 mg: Folgende Gründe führten zum Medikationsabbruch nach Randomisierung Widerruf der Einwilligungserklärung: n=9 Unerwünschtes Ereignis: n=16 Unbefriedigender therap. Effekt: n=5 Abnormer Laborwert: n=7 Administrative Probleme: n=2 Protokollverletzung: n=1 Abnormes Testergebnis: n=4 INF β -1a i. m.: Folgende Gründe führten zum Medikationsabbruch nach Randomisierung Widerruf der Einwilligungserklärung: n=16 Unerwünschtes Ereignis: n=12 Unbefriedigender therap. Effekt: n=7 Abnormer Laborwert: n=3 Erschien nicht mehr zu den Visiten: n=4 Administrative Probleme: n=3 Protokollverletzung: n=2 Abnormes Testergebnis: n=4
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der	FPFV: 30. Mai 2006 LPLV: 11. November 2008

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

	Nachbeobachtung	
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie endet planmäßig nach der letzten Visite des letzten in der Studie behandelten Patienten. Das Rekrutierungsziel war erreicht.
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

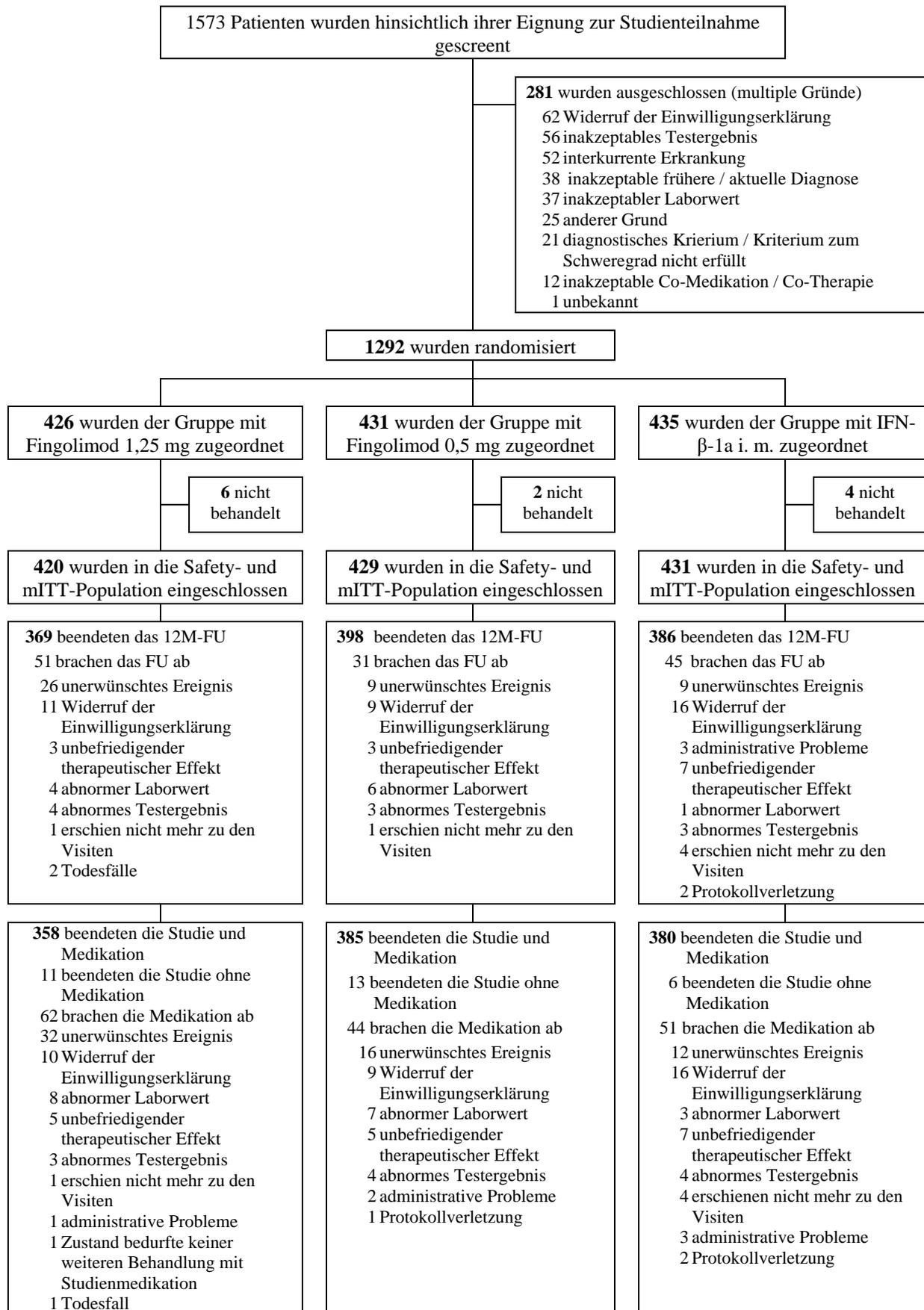


Tabelle 4-89 (Anhang): Studiendesign und -methodik für Studie Bornstein et al., 1987

Item ^a	Charakteristikum	Studieninformation
-------------------	------------------	--------------------

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Primärziel: Anteil schubfreier Patienten Sekundärziele: Schubhäufigkeit, Veränderung des Behinderungsgrades (Kurtzke Score) zur Baseline, Zeitspanne bis Verschlechterung.
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Pilotstudie, doppelblind, randomisiert, Placebo-kontrolliert. Matched-Pair-Randomisierungsverhältnis 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine relevanten Änderungen der Methodik nach Studienbeginn.
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: 1. diagnostizierte MS 2. Alter zwischen 20-35 Jahre 3. überdurchschnittliche Schubrate von mindestens zwei diagnostizierten und dokumentierten Schüben innerhalb von 2 Jahren vor Studienbeginn. 4. Kurtzke (EDSS) - Skala 0-6 5. Emotional stabil laut psychosozialen Gutachten. Ausschlusskriterien*: 932 Fragebögen von Freiwilligen wurden ausgewertet, von diesen wurden 140 näher neurologisch und psychosozial untersucht. 90 Patienten wurden ausgeschlossen wegen folgenden Gründen: 23 Alter 21 zu wenig Schübe im Vorfeld 19 fehlende Dokumentation 15 psychosozial unpassend 8 Übergang zur chronischen Form 3 Distanz zur Klinik 1 Schwangerschaft
4b	Studienorganisation und Ort der Studiendurchführung	USA (1 Zentrum)
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Prüfsubstanz: COP 1 (Glatirameracetat) in bakteriostatischer Kochsalzlösung (20 mg / ml) Kontrollsubstanz: Kochsalzlösung (1 ml) Beide Substanzen liegen in Ampullenform vor und sind bis zum Gebrauch bei -20° C eingefroren. Jeder Patient bekommt eine Monatspackung

Item ^a	Charakteristikum	Studieninformation
		mit 32 Ampullen. Nach Auftauen der Tagesration selbstständige Injektion durch Patient (nach entsprechender Schulung).
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 Zielkriterium: Anteil der schubfreien Patienten Sekundäre Zielkriterien: Schubhäufigkeit. Veränderungen des Behinderungsgrades (Kurtzke (EDSS) - Skala), auf Baseline bezogen. Zeitspanne bis Verschlechterung.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Änderung der Zielkriterien.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Pilotstudie, keine Angaben zur Bestimmung der Fallzahlen.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	N/A
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Keine Angaben *
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Es wurden Matched-Pairs aufgrund folgender Angaben gebildet: <ul style="list-style-type: none"> - Geschlecht - Anzahl der Schübe innerhalb von 2 Jahren vor Studienbeginn - Grad der Behinderung nach Kurtzke-Skala: die 3 Strata 0-2, 3-4 und 5-6 Die Randomisierungs-Zuweisung des ersten Patienten eines Paares bestimmte die des zweiten.
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	Randomisierung 1:1 für jedes Matched-Pair durch unverblindetes Zentrumspersonal. Dieses war unter Anleitung eines Statistikers für die verblindete Kennzeichnung der Medikation verantwortlich.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Patienten wurden nach Stratifizierung einem Matched-Pair zugewiesen. Randomisierung 1:1 für jedes Matched-Pair durch unverblindetes Zentrumspersonal.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung	a) Die Patienten waren verblindet. b) Zentrumspersonal, das für die Zubereitung, Kennzeichnung und Verteilung der Ampullen verantwortlich war, war unter Anleitung des für die Randomisierung der Patienten

Item ^a	Charakteristikum	Studieninformation
	vorgenommen?	<p>verantwortlichen Statistikers unverblindet.</p> <p>c) Der Neurologe, der neurologische Untersuchungen und Status-Befunde erhob, war verblindet.</p> <p>Sämtliche Diskussionen über Nebenwirkungen o.ä. zwischen Patient und untersuchendem Neurologen wurden vermieden. Die Patienten meldeten diese nur dem unverblindeten Studienkoordinator.</p>
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Keine Angabe
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Die Matched-Pairs-Analyse hinsichtlich der Unterschiede zwischen den Behandlungsarmen wurde nach McNemar getestet.</p> <p>Ein zweiseitiger exakter Fischer-Test wurde bei den 2:2 Kontingenztabelle, der Chi-Quadrat-Test für die 2:3 Kontingenztabelle zur Schubhäufigkeit angewandt.</p> <p>Kurven zur Darstellung der Dauer bis zur Progression (definiert als Anstieg der Kurztk-Skala um mindestens einen Punkt und für mindestens 3 Monate anhaltend) wurden mit Life-Table-Methoden errechnet.</p> <p>Ein Log-Rank-Test wurde zum Vergleich der Überlebenskurven für jede Behandlungsgruppe durchgeführt, ebenso zum festen Zeitpunkt nach 24 Monate.</p> <p>Multiple logistische Regressionsanalysen wurden für die Auswirkungen auf das Behandlungsergebnis im Bezug auf Geschlecht, Erkrankungsdauer, vorangegangene Schubraten, Behinderungsgrad bei Studieneintritt und verschiedene Kombinationen dieser Variablen durchgeführt.</p> <p>Die ODDs Ratio wurde aus den Regressionskoeffizienten bestimmt.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Geplante Subgruppenanalysen bezüglich des Behinderungsgrades.</p> <p>Baseline-Charakteristika der Studienpopulation wurden mittels eines zweiseitigen t-Tests für stetige Variablen und eines Chi-Quadrat-Tests mit Yates-Korrekturen für diskrete Variablen durchgeführt.</p> <p>Unterschiede in den Nebenwirkungen wurden mit dem Chi-Quadrat-Test bewertet.</p>
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	

Item ^a	Charakteristikum	Studieninformation
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	a) 50 Patienten wurden randomisiert, 48 wurden 24 Matched-Pairs zugewiesen, 2 Patienten wurden jeweils einer Behandlungsgruppe zugewiesen. b) 50 Patienten erhielten die Behandlung.* c) 22 Matched-Pairs plus 4 Patienten in unpaariger Analyse
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Placebo-Gruppe: 2 Patienten wurden wegen unbrauchbaren Daten ausgeschlossen. Für den Abbruch der Studie waren psychologische Gründe ausschlaggebend.
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Keine Angaben *
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde wie geplant abgeschlossen.

a: nach CONSORT 2010.

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

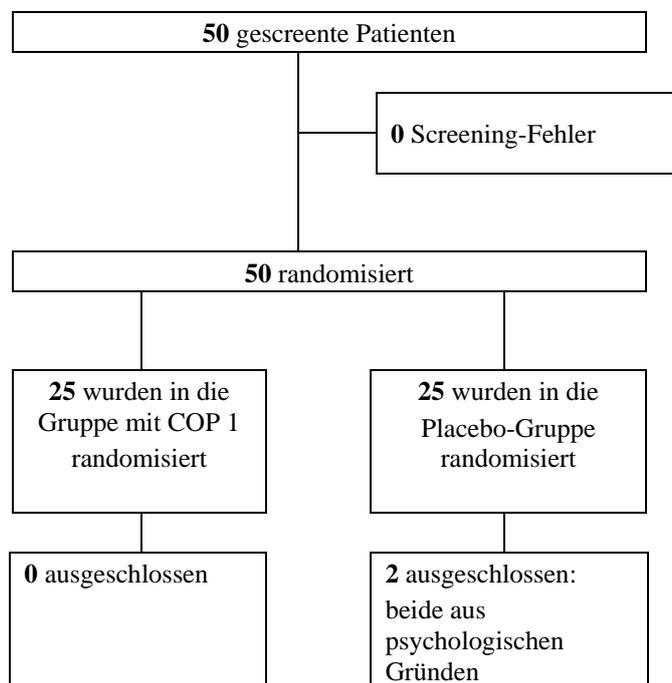


Tabelle 4-90 (Anhang): Studiendesign und -methodik für Studie BECOME

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Primärziel Anzahl der kombinierten aktiven Läsionen (CAL) pro Patient pro Scan Sekundärziel Anzahl neuer Läsionen und klinischer Schübe
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisiert, parallel, offen, auswerterverblindet, Phase IV Randomisierungsverhältnis: 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Angaben *
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien - Patienten zwischen 18 und 55 Jahren - vorhandene Einwilligungserklärung - Vorliegen einer demyelinisierenden Erkrankung: entweder - RRMS nach überarbeiteten McDonald Kriterien plus Nachweis einer akuten Krankheitsaktivität, nachgewiesen durch die Entwicklung eines oder mehrerer klinischer und/oder MRT –Auffälligkeiten während der vergangenen 6 Monate oder - CIS (Clinically isolated syndrome), charakteristisch für eine ZNS-Demyelinisierung (optische Neuritis oder transverse Myelitis oder Hirnstamm-Läsion) bestätigt durch Untersuchung mit Beginn innerhalb der vergangenen 6 Monate plus Nachweis einer zeitlichen und räumlichen Ausdehnung. Der Nachweis einer räumlichen Ausdehnung erforderte zwei oder mehr Hirn-Läsionen im MRT von denen mindestens eine ovoid und/oder periventrikulär sein musste. Der Nachweis einer zeitlichen Ausdehnung für akute CIS (≤ 1 Monat) erforderte ≥ 1 nicht-anreichernde Hirn-Läsion und für nicht-akute CIS (> 1 Monat) ≥ 1 anreichernde Hirn-Läsion. - EDSS von 0 bis einschließlich 5,5 an Baseline - Frauen im gebärfähigen Alter mussten eine adäquate Form der Empfängnisverhütung anwenden - Vorliegen eines negativen Schwangerschaftstest (Bluttest bei Screening und Urintest an Baseline)

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> - Adäquate Funktion des Knochenmarks, der Nieren und der Leber für alle Patienten durch Laborwerte bei Screening bestätigt Ausschlusskriterien - Ausbruch eines Schubes zwischen Screening und Tag 1 der Studie - Vorgeschichte zugrundeliegender Erkrankungen, die das ZNS betreffen oder mit den MRT-Ergebnissen oder anderen Studienuntersuchungen interferieren könnten - übliche Metallgegenstände oder Fremdkörper, die bei MRTs kontraindiziert sind - Inkompatible Größe oder Gewicht für die speziell dafür ausgelegte 3T Kopf-MRT Einheit - schwangere oder stillende Frauen - bekannte Allergie oder Überempfindlichkeit gegen Gadolinium-Chelate, Albumine, Interferone, GA oder Mannitol - unkontrollierte oder klinisch signifikante Herzerkrankung - Vorgeschichte eines instabilen medizinischen Zustandes, der als klinisch signifikant erachtet werden könnte - Unverträglichkeit oder jede Kontraindikation gegen Acetaminophen, Ibuprofen oder Steroide - Unfähigkeit die Protokollanforderungen zu erfüllen, im Ermessen des Arztes - Teilnahme an einer klinischen Prüfung innerhalb der vergangenen 6 Monate - gegenwärtige Abhängigkeit - Magen- oder Zwölffingerdarmgeschwür - Unfähigkeit subkutane Injektionen zu applizieren, entweder selbst oder durch einen Betreuer - medizinische, psychiatrische oder andere Zustände, die die Fähigkeit des Patienten beeinflussen könnten, die Studieninformation zu verstehen, die Einwilligung zu geben, das Studienprotokoll zu befolgen oder die Studie zu beenden - Klaustrophobie - unkontrollierte Kopfbewegungen - Behandlung mit irgendeinem Interferon oder GA - vorangegangene Bestrahlung lymphatischer Organe des gesamten Körpers, vorangegangene Verwendung monoklonaler Antikörper (z. B. anti-CD25, anti-CD52, anti VLA-4, anti-CD20), Mitoxantron, Cyclophosphamid, Cladribin, Azathioprin, Mycophenolat, IVIG oder Cyclosporin A;

Item ^a	Charakteristikum	Studieninformation
		- Kortikosteroide innerhalb von 21 Tagen vor Studienbeginn
4b	Studienorganisation und Ort der Studiendurchführung	2 Zentren in USA (New Jersey)
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Prüfsubstanzen - IFN β -1b (Betaseron®, Bayer HealthCare Pharmaceuticals, Montville, NJ), 250 μ g alle zwei Tage - Glatirameracetat (Copaxone®, Teva-Marion Partners, Kansas City, MO), 20 mg täglich
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 Zielkriterium Anzahl der kombinierten aktiven Läsionen (CAL) pro Patient pro Scan (monatliche MRTs) während des ersten Jahres, einschließlich aller anreichernden und nicht-anreichernden neuen T2/*Fluid-attenuated inversion recovery* (FLAIR) Läsionen Sekundäre Zielkriterien - Anzahl neuer Läsionen (NL) pro Patient in Jahr 1 und während Jahr 2 - Anzahl der klinischen Schübe - Anzahl der bezüglich CAL erkrankungsfreien Patienten über 2 Jahre Ein Schub ist definiert als: Alle neuen oder sich verschlechternden Symptome, die \geq 24 Stunden andauern und nicht durch Fieber oder eine Infektion erklärt werden können, wurden als subjektive Schübe erachtet. Subjektive Schübe, die durch einen verblindeten untersuchenden Neurologen bestätigt wurden, unter Verwendung entweder der „Scripps Neurological Rating Scale“ (SNRS) oder dem „Expanded Disability Status Scale“ (EDSS), wurden als objektive Schübe erachtet. Mindestens eine der folgenden Veränderungen im Vergleich zu Baseline war zur Bestätigung des Schubes erforderlich: 1) Erhöhung des Gesamt-EDSS um \geq 0,5 Punkte; 2) Erhöhung des EDSS-Wertes für ein System um \geq 2 Punkte; 3) Erhöhung des Wertes von 2 oder mehr EDSS-Systemen um \geq 1 Punkt; 4) Erniedrigung des SNRS-Wertes um \geq 7 Punkte. Allen Patienten, bei denen ein Schub durch einen verblindeten Neurologen bestätigt wurde, wurde eine Behandlung mit 1 g Methylprednisolon täglich für 5 Tage angeboten.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Angabe*
7	Fallzahl	

Item ^a	Charakteristikum	Studieninformation
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahl wurde ausgewählt um die Überlegenheit zu zeigen, unter Annahme einer 2/3 Wahrscheinlichkeit, dass ein zufällig ausgewählter IFN β -1b behandelter Patient weniger CAL hat, als ein zufällig ausgewählter GA Patient, basierend auf vorhergehenden monatlichen Hirn-MRT-Studien. Eine Fallzahl von 40 Patienten pro Arm ergab eine 74%-ige Power bei einem zwei-seitigen Test mit $\alpha = 0,05$.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Da die initiale Interimsanalyse keine Sicherheitsbedenken der Anwendung von monatlichen dreifach-Dosen Gadolinium hervorbrachte, wurde allen Patienten in der Studie die Option angeboten, zusätzliche monatliche MRT-Scans für ein zweites Behandlungsjahr zu erhalten.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Keine Angabe*
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Stratifizierung nach Prüfzentrum und je nach Vorhandensein einer Anreicherung beim Screening-MRT.
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	Keine Angabe*
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Randomisierung wurde pro Prüfzentrum (Newark oder Teaneck) und nach Vorhandensein einer Anreicherung beim Screening-MRT durchgeführt.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) Die Patienten konnten, aufgrund der charakteristischen Injektionsreaktionen auf IFN β -1b oder GA, nicht verblindet werden. b) diejenigen, die die Intervention/ Behandlung durchführten waren nicht verblindet c) Die MRT-Auswerter waren gegenüber der Behandlung verblindet.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht nötig, da die Patienten nicht verblindet waren.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Wirksamkeitsanalysen wurden nach Intent-to-treat durchgeführt. Fisher-Exact-Test wurde angewendet, um die Häufigkeiten der Baseline Charakteristika, mit den Anteilen an Patienten, deren Scans läSIONsfrei blieben, zu vergleichen. Rangsummen-Tests wurden verwendet, um die Verteilung von kontinuierlichen und ordinalen

Item ^a	Charakteristikum	Studieninformation
		<p>Charakteristika zu vergleichen. Der primäre Wirksamkeitsendpunkt war CAL pro Scan im ersten Jahr, mit Patient als Analyseeinheit. Die Verteilung der gezählten Endpunktläsionen in den Studienarmen wurde mittels Rangsummen-Test verglichen, stratifiziert nach An- oder Abwesenheit von anreichernden Läsionen bei Screening oder Baseline unter Verwendung von „van Elteren’s modifiziertem Ridit-Score“. Die Volumina der anreichernden Läsionen wurden zusammengefasst als Gesamtvolumen der Läsionen pro Scan pro Patient und die Behandlungen wurden mittels Rangsummentest verglichen, hierbei wurden Patienten ohne Läsionen ausgeschlossen.</p> <p>Jede Null-Hypothese bzgl. sekundärer Endpunkte wurde mit zweiseitigem $\alpha = 0,05$ getestet, ohne Adjustierung der Alphafehler-Kumulierung. SAS statistische Software Version 9.1 wurde verwendet.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Die jährliche Schubrate (ARR) wurde berechnet als Gesamtzahl an Schüben geteilt durch die Gesamtzeit in der Studie. ARR nach Behandlungsgruppen wurde verglichen unter Verwendung des Poisson Regressions-Modells. Die Zeit bis zum ersten Schub wurde verglichen unter Verwendung einer Log-Rank-Überlebensanalyse.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<p>168 Patienten gescreent</p> <p>a) 75 randomisiert</p> <p>b) 36 Patienten erhielten IFNβ-1b und 39 Patienten erhielten GA</p> <p>c) 36 Patienten (IFNβ-1b) und 39 Patienten (GA) wurden analysiert</p>
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>IFNβ-1b</p> <p>Von 36 randomisierten Patienten waren 7 ohne aktuelle Verlaufskontrolle</p> <ul style="list-style-type: none"> - 2 Schwangerschaften - 1 Brustkrebs - 1 Schlaganfall - 3 ohne aktuelle Verlaufskontrolle <p>4 haben die Behandlung abgebrochen</p> <ul style="list-style-type: none"> - 1 wegen rezidivierender Hautnekrose - 3 wegen Behandlungsausschluss durch den behandelnden Neurologen

Item ^a	Charakteristikum	Studieninformation
		GA Von 39 randomisierten Patienten waren 4 ohne aktuelle Verlaufskontrolle - 1 zu lange Studiendauer - 1 umgezogen - 2 Stress 4 haben die Behandlung abgebrochen - 4 wegen Behandlungsausschluss durch den behandelnden Neurologen
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Studie begann im Januar 2003 und endete im Januar 2006.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde nach 1 Jahr um 1 weiteres Jahr verlängert und nach 2 Jahren wie geplant abgeschlossen.
a: nach CONSORT 2010.		

* Keine Angaben hierzu in der Publikation.

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

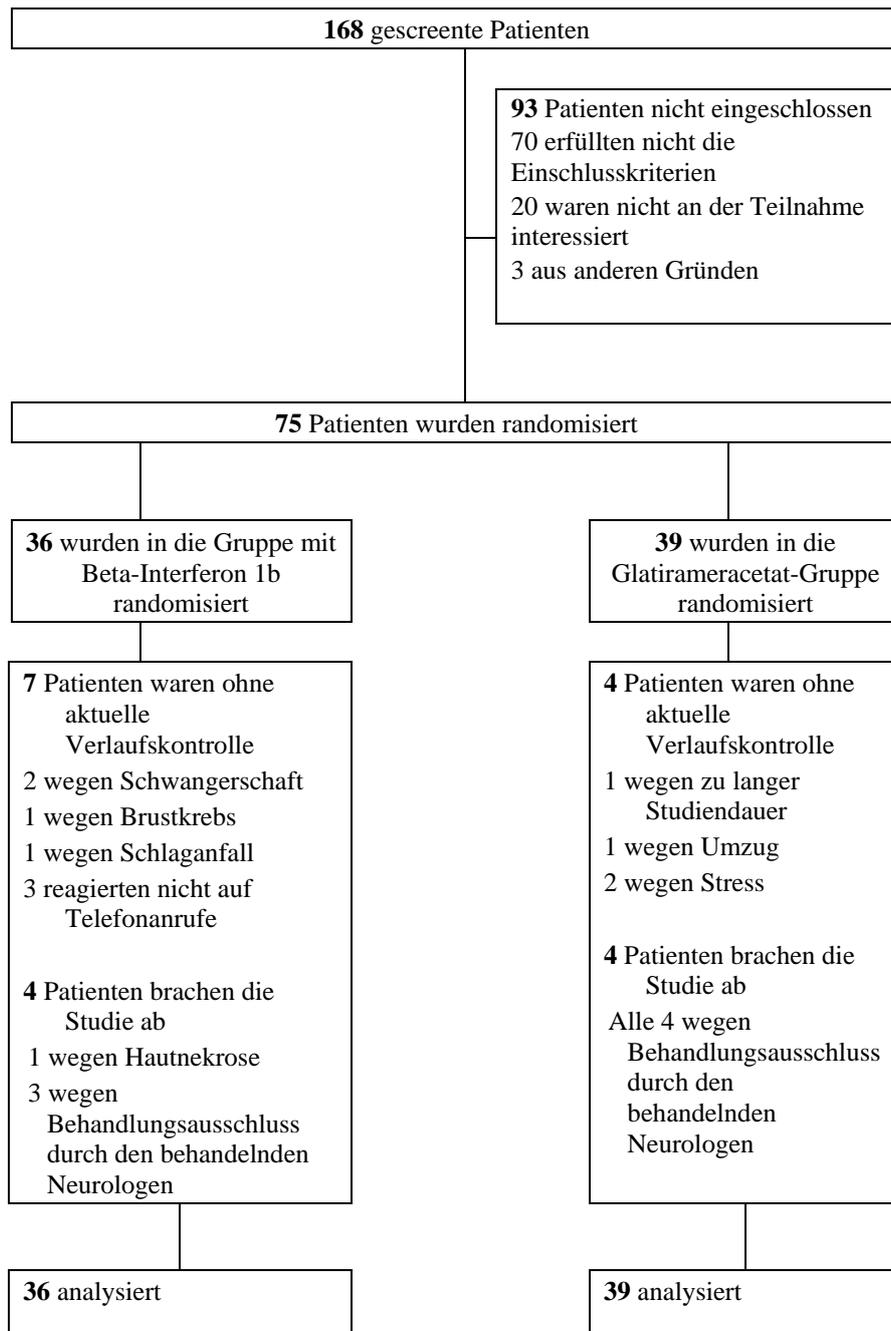


Tabelle 4-91 (Anhang): Studiendesign und -methodik für Studie INCOMIN

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	<p>Die INCOMIN-Studie wurde geplant, um Beta-Interferon-1b (250 µg, alle zwei Tage) mit Beta-Interferon-1a (30 µg, einmal wöchentlich) hinsichtlich des klinischen Nutzens und des Nutzens einer Magnetresonanztomographie zu vergleichen.</p> <p>Primäre Zielparameter:</p> <ul style="list-style-type: none"> • Anteil schubfreier Patienten; • Anteil der Patienten, frei von neuen Protonendichte/T2-Läsionen; <p>Sekundäre Zielparameter:</p> <ul style="list-style-type: none"> • Jährliche (behandelte) Schubrate; • Anteil der Patienten ohne fortschreitende Behinderung; • Dauer bis zum Fortschreiten der Behinderung; • Anteil der Patienten, frei von Gadolinium-angereicherten Läsionen und ohne „MRT-Aktivität“ (definiert als neues Vorkommen einer Protonendichte/hyperintensives T2 oder vermehrter Läsionen während der Studie);
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	<p>prospektive, randomisierte, parallele multi-zentrische Studie;</p> <p>Randomisierung 1:1</p>
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Angaben*
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <ul style="list-style-type: none"> • Erwachsene (18 – 50 Jahre) mit RRMS; • mindestens zwei Schübe in den letzten zwei Jahren; • kein Schub (und keine Corticosteroid-Behandlung) innerhalb 30 Tage vor Studieneintritt • EDSS zwischen 1,0 und 3,5; <p>Ausschlusskriterien:</p> <ul style="list-style-type: none"> • vorherige systemische Behandlung mit Beta-Interferon oder Behandlung mit anderen immunsuppressiven oder immuno-modulierenden Medikamenten (außer Corticosteroide); • Schwangerschaft, Stillzeit oder fehlende Bereitschaft, eine akzeptable Empfängnisverhütung einzuhalten; • Größere Depressionen oder Selbstmordversuche

Item ^a	Charakteristikum	Studieninformation
		in der medizinischen Vorgeschichte; <ul style="list-style-type: none"> • Klinisch signifikante Herz-, Leber-, Nieren- oder Knochenmarkserkrankung;
4b	Studienorganisation und Ort der Studiendurchführung	15 Zentren in Italien Die Studie wurde an der Universität von Turin durch Professor Durelli koordiniert.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Prüfsubstanzen: <ul style="list-style-type: none"> • Beta-Interferon-1b (250 µg, s. c., alle zwei Tage); • Beta-Interferon-1a (30 µg, i. m., einmal wöchentlich); Die Patienten wurden zufällig eingeteilt, um sich die jeweilige Prüfsubstanz selber zu verabreichen.
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	Primäre Zielparameter: <ul style="list-style-type: none"> • Anteil schubfreier Patienten innerhalb 24 Monate der Studie; • Anteil der Patienten, frei von neuen Protonendichte/T2-Läsionen; Sekundäre Zielparameter: <ul style="list-style-type: none"> • Jährliche (behandelte) Schubrate; • Anteil der Patienten ohne fortschreitende Behinderung; • Dauer bis zum Fortschreiten der Behinderung; • Anteil der Patienten, frei von Gadolinium-angereicherten Läsionen und frei von „MRT-Aktivität“; Reguläre Visiten waren für die Tage 1, 3, 5, 15 und 30 und danach alle 3 Monate geplant. Ungeplante Visiten zur Beurteilung und Behandlung von Schüben oder für andere medizinische Ereignisse ohne MS-Bezug wurden je nach Notwendigkeit durchgeführt und dokumentiert. Magnetresonanztomographie-Scans wurden zentral durch Prüfer an der Universität von Turin im MRT-Analyse-Zentrum beurteilt.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Angaben*
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahlen wurden auf der Basis publizierter Ergebnisse der pivotalen Studien mit Beta-Interferon 1b (alle zwei Tage) und Beta-Interferon-1a (einmal wöchentlich) ermittelt. Unter der Annahme eines erwarteten Patienten-Anteils mit Schüben von 63 % in der Beta-Interferon-1a-Gruppe benötigt man für zwei Gruppen gleicher Größe mit einer statistischen Power von 80 % und einem Konfidenzintervall von 95 % 236 Patienten, um ein relatives Schub-Risiko von 0,7 zu entdecken.
7b	Falls notwendig, Beschreibung von	Zwischenanalysen wurden nicht geplant.

Item ^a	Charakteristikum	Studieninformation
	Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Randomisierungs-Codes wurden durch ein unabhängiges Team von Statistikern (Department of Public Health, Medical School der Universität von Turin) den Behandlungen zugeordnet, in Unkenntnis der klinischen Eigenschaften des Patienten.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierung erfolgte nach computer-generierten Zufalls-Sequenzen von Ziffern, die sich für jedes Zentrum und für jedes Geschlecht unterschieden, um eine Zentrums- und Geschlechts-Stratifikation zu erreichen. Alle in die Studie eingeschlossenen Patienten wurden nachfolgend randomisiert. Es wurden keine Maßnahmen ergriffen, um die Fallzahl in den zwei Randomisierungsgruppen, wie z.B. Block- oder Austausch-Randomisierung, auszugleichen. Dadurch wurde eine mögliche nicht-zufällige Manipulation der Patienten-Zuteilung zu jeder Behandlungsgruppe vermieden.
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	Randomisierung wurde zentral durch das Randomisierungs-Zentrum durchgeführt.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Das Screening erfolgte an den jeweiligen Zentren in Italien. Koordinierende Neurologen wiesen den in die Studie eingeschlossenen Patienten einen Randomisierungs-Code zu, mit detaillierter Angabe ihrer fortlaufenden Einschlussnummer, Zentrumscode und Geschlecht.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) Die Patienten waren nicht verblindet. b) Die Patienten verabreichten sich die Medikation selbst. c) Die Ärzte, die die Beurteilungen vornahmen, waren nicht verblindet. Die Ergebnisse der Magnetresonanztomographie stammten aus einer komplett verblindeten zentralen Analyse im MRT-Analysezentrum der Universität von Turin, in Unkenntnis der Behandlungszuteilung und der klinischen Eigenschaften des Patienten. Selektionstendenzen wurden durch eine zentralisierte Randomisierung mittels einer verborgenen Behandlungszuteilung gesteuert. Untersuchungstendenzen waren vermutlich aufgrund konsistenter klinischer Ergebnisse marginal.

Item ^a	Charakteristikum	Studieninformation
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht nötig, da die Patienten nicht verblindet waren.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Analysen wurden nach Intent-to-Treat durchgeführt. Alle Daten wurden unter Verwendung der gesamten 2-Jahresdaten der Patienten, die die Behandlung absolvierten oder die die Behandlung abbrachen, eingerechnet. Die Daten der Patienten, die zur Nachuntersuchung nicht antraten, wurden in der Analyse bis zu dem Punkt behalten, an dem sie verloren gingen; für nachfolgende Analysen wurden diese Patienten als „schlecht“ bezüglich des Ergebnisses betrachtet.</p> <p>Der Chi-Quadrat-Test, oder wenn nötig der Fisher's Exakt-Test, wurden für dichotome Endpunkte und parametrische oder nicht-parametrische Tests wurden für stetige Endpunkte, je nach der Verteilung der Variablen, verwendet, um Signifikanzen zu testen. Relatives Risiko, absolute Risikoverminderung und die Anzahl der Patienten, die behandelt werden mussten, um einen schlechtes Ergebnis zu vermeiden, wurden für dichotome Endpunkte, zusammen mit dem 95 %-Konfidenzintervall, und für die standardisierten Mittelwertunterschiede (SMU) für stetige Endpunkte, zusammen mit dem 95 %-Konfidenzintervall, geschätzt. Für den zeitlichen Verlauf bis zum Ereignis wurde die „Life-Table-Methode“ verwendet und die Signifikanz wurde mit dem „Log-Rank-Test“ getestet. Die Daten wurden als direkter Vergleich der zwei Behandlungen analysiert und hinsichtlich des Schweregrades der Erkrankung zur Baseline stratifiziert. Die Schubrate in den zwei Jahren vor der Studie, der EDSS-Score zur Baseline, die Zahl der Protonendichte/T2 hyperintenser oder Gadolinium-angereicherter Läsionen wurden als Indikatoren des klinischen oder auf MRTs basierenden Erkrankungs-Schweregrades angesehen.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Keine Angaben*
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die <ul style="list-style-type: none"> a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden 	<p>Beta-Interferon-1b (250 µg, s. c., alle zwei Tage):</p> <ul style="list-style-type: none"> a) 96 Patienten / MRT-Teilnahme: 76 Patienten; b) 96 Patienten / MRT-Teilnahme: 76 Patienten; c) 96 Patienten / MRT-Teilnahme: 76 Patienten; <p>Beta-Interferon-1a (30 µg, i. m., einmal wöchentlich):</p> <ul style="list-style-type: none"> a) 92 Patienten / MRT-Teilnahme: 73 Patienten; b) 92 Patienten / MRT-Teilnahme: 73 Patienten;

Item ^a	Charakteristikum	Studieninformation
		c) 92 Patienten / MRT-Teilnahme: 73 Patienten;
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Beta-Interferon-1b (250 µg, s. c., alle zwei Tage): Abbrüche insgesamt: n = 11 5 wegen unerwünschter Ereignisse 3 wegen andauernder Erkrankungs-Aktivität oder -Progression 1 wegen Schwangerschaft 2 ohne aktuelle Verlaufskontrolle</p> <p>Beta-Interferon-1a (30 µg, i. m., einmal wöchentlich): Abbrüche insgesamt: n = 19 1 wegen einem unerwünschten Ereignis 10 wegen andauernder Erkrankungs-Aktivität oder -Progression 2 wegen Behandlungsverweigerung 2 wegen Schwangerschaft 4 ohne aktuelle Verlaufskontrolle</p>
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Rekrutierungsphase der Studie begann im Oktober 1997 und endete im Juni 1999. Im Juni 2001 wurde die Studie beendet.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde wie geplant abgeschlossen.
a: nach CONSORT 2010.		

* Keine Angaben hierzu in der Publikation.

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

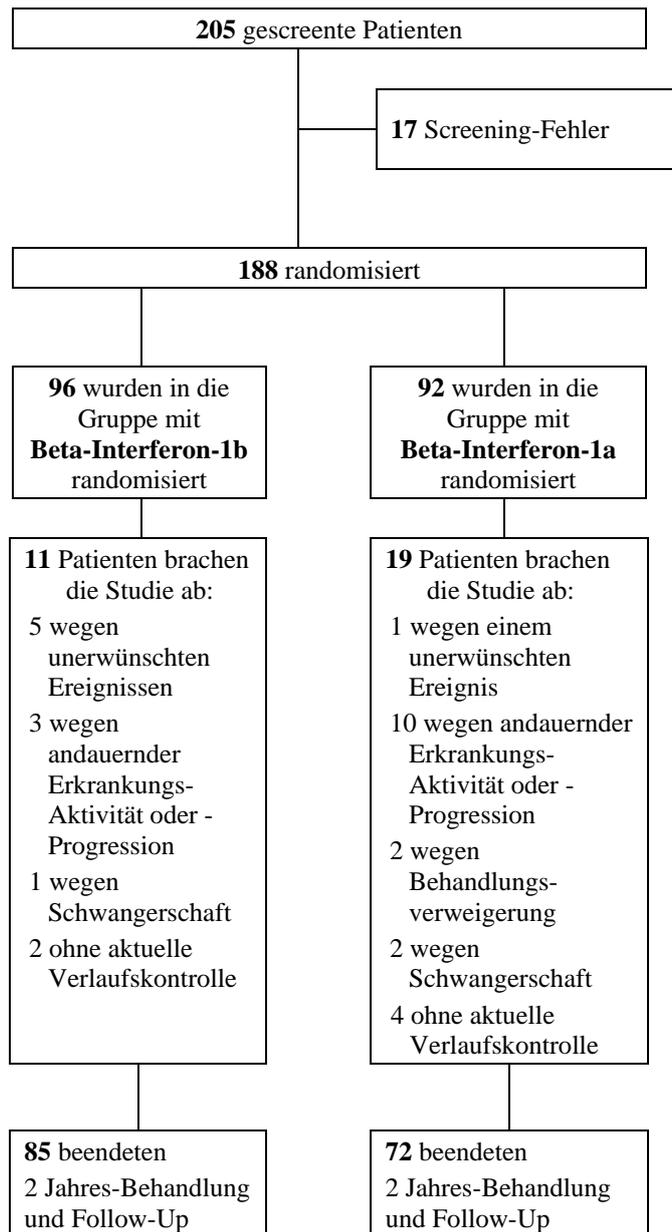


Tabelle 4-92 (Anhang): Studiendesign und -methodik für Studie Etemadifar et al., 2006

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Vergleich der relativen Wirksamkeiten von Betaferon, Avonex und Rebif in der Behandlung von RRMS. Primärziel: Bewertung der relativen Wirksamkeiten der drei Interferon-Produkte unter Verwendung der klinischen Standard-Dosierungen im Hinblick auf klinische Schübe und EDSS.
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Einfachblind (auswerterverblindet), randomisiert, parallel Randomisierungsverhältnis 1:1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine relevanten Änderungen
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: <ol style="list-style-type: none"> 1. Männer und Frauen (15-50 Jahre) 2. Klinisch und labortechnisch nachgewiesene RRMS 3. Mindestens 2 vom Neurologen dokumentiert Schübe in den letzten zwei Jahren vor Studienbeginn 4. EDSS ≤ 5 Ausschlusskriterien: <ol style="list-style-type: none"> 1. Schwere Allergie oder anaphylaktische Reaktion auf Interferon oder anderen Bestandteilen der Medikation. 2. Keine Anzeichen von neurologischen, psychiatrischen, kardiologischen, endokrinologischen, hämatologischen, hepatischen, renalen Erkrankungen. 3. Keine Anzeichen von bestehenden Krebserkrankungen, Auto-Immun-Erkrankungen oder anderen chronischen Erkrankungen. 4. Keine Anzeichen von unkontrollierten Krampfanfällen, Suizidgedanken oder Episoden schwerer Depression in den letzten 3 Monaten vor Studienbeginn. 5. Schwangere oder stillende Frauen. Gebärfähige Frauen müssen eine klinisch akzeptierte Methode der Empfängnisverhütung anwenden.
4b	Studienorganisation und Ort der Studiendurchführung	2 Zentren im Iran
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Prüfsubstanzen: • Betaferon (IFN-β-1b, 250 µg s. c., alle 2 Tage)

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Avonex (IFN-β-1a 30 µg, i. m., 1 x pro Woche) • Rebif (IFN-β-1a 44 µg, s. c. 3 x pro Woche) <p>Alle Patienten verabreichten sich die Medikation selbst.</p>
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primäre Zielkriterien: Anzahl der Schübe; EDSS-Veränderungen; Akuter Schub ist definiert als das Auftreten eines neuen neurologischen Symptoms oder die ernste Verschlimmerung eines vorbestehenden Symptoms, das mindestens 24 Stunden anhält und eine Verschlechterung des EDSS um einen Punkt nach sich zieht. Die Patienten wurden im Monat 6, 12 und 24 nach Therapiestart durch einen qualifizierten Neurologen auf Nebenwirkungen und Compliance bewertet.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Änderungen
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Keine Angaben [*]
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es wurden keine Zwischenanalysen durchgeführt. [*]
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Keine Angaben [*]
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Keine Angaben [*]
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	Keine Angaben [*]
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Keine Angaben [*]
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	<p>a) Die Patienten waren nicht verblindet. b) Die Patienten verabreichten sich die Medikation selbst. c) Die Ärzte, die die Beurteilungen vornahmen, waren verblindet.</p>
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht nötig, da die Patienten unverblindet waren und sich die Medikation selbst

Item ^a	Charakteristikum	Studieninformation
		verabreichten.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Für Vergleiche zwischen den 3 Behandlungsgruppen wurde die univariate Varianzanalyse (ANOVA) durchgeführt. Zum zeitlichen Verlauf wurde die Varianzanalyse (ANOVA) mit wiederholten Messungen angewandt. Für Vergleiche zu den verschiedenen Zeitpunkten innerhalb jeder Gruppe wurde der paarweise Student t-Test angewandt. Prozentuale Vergleiche wurden mit dem Chi-Quadrat oder exakten Fischer-Test getestet. Alle statistischen Tests waren zweiseitig. Die Analysen wurden auf einem PC mit Hilfe der SPSS-Analyse-Software für Windows durchgeführt.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Es wurden keine Subgruppenanalysen durchgeführt.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Betaferon (IFN-β-1b, 250 µg, s. c.) a) 30 Patienten b) 30 Patienten c) 30 Patienten Avonex (IFN-β-1a, 30 µg, i. m.) a) 30 Patienten b) 30 Patienten c) 30 Patienten Rebif (IFN-β-1a, 44 µg, s. c.) a) 30 Patienten b) 30 Patienten c) 30 Patienten
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Keine Ausschlüsse nach Randomisierung. Alle Patienten waren bis zum Schluss in der Studie.
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Patienten wurden zwischen September 2002 und September 2004 eingeschlossen.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde planmäßig beendet.
a: nach CONSORT 2010.		

* Keine Angaben hierzu in der Publikation.

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

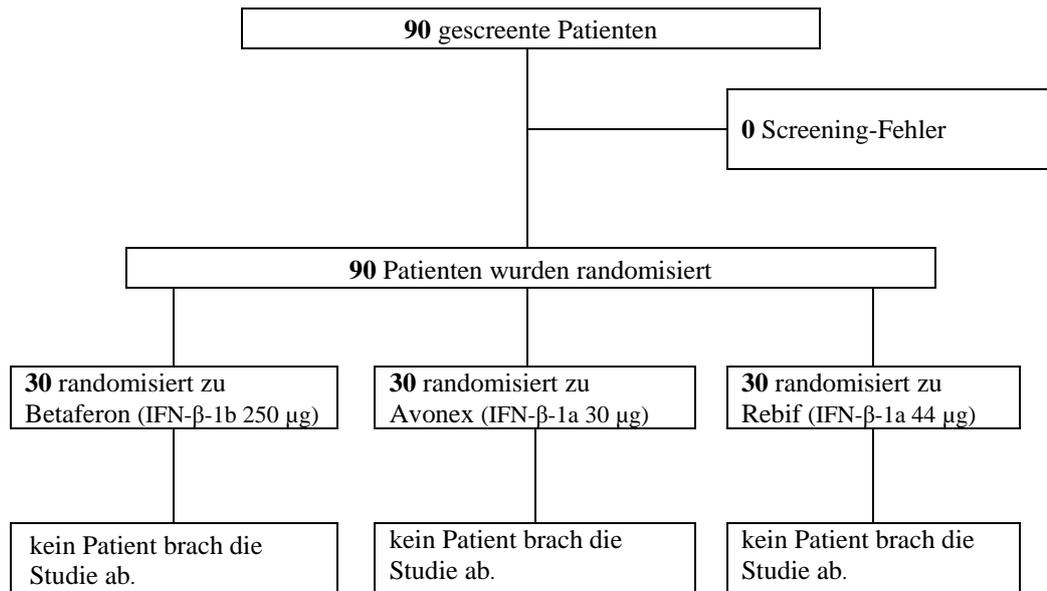


Tabelle 4-93 (Anhang): Studiendesign und -methodik für Studie IFNB-MS-Study Group

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	<p>Primärziel Unterschiede in der Schubrate und Anteil an Patienten, die schubfrei bleiben;</p> <p>Sekundärziele</p> <ul style="list-style-type: none"> - Zeit bis zum ersten Schub; - Dauer und Schwere des Schubs; - Veränderung in EDSS und NRS-Werten zu Baseline; - Quantitative Messung der Krankheitslast durch jährliches MRT; - Messung der Krankheitsaktivität;
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Multizentrisch, randomisiert, doppelblind, Placebo-kontrolliert, Pilotstudie zur Dosisfindung; Randomisierungsverhältnis 1:1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Angaben*
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien</p> <ul style="list-style-type: none"> • Alter zwischen 18 und 50 Jahre; • EDSS zwischen 0 und 5,5; • mindestens zwei Schübe in den vergangenen zwei Jahren; <p>Ein Schub ist definiert als Auftreten eines neuen Symptoms oder Verschlechterung eines bestehenden Symptoms, bedingt durch MS; begleitet von einer entsprechenden neurologischen Auffälligkeit; mindestens 24 Stunden andauernd ohne Fieber; gefolgt von einer Stabilisierung oder Verbesserung für mindestens 30 Tage.</p> <ul style="list-style-type: none"> • klinisch stabil für mindestens 30 Tage vor Studieneintritt; <p>Ausschlusskriterien</p> <ul style="list-style-type: none"> • vorherige Gabe von ACTH oder Prednison innerhalb von 30 Tagen; • vorangegangene Behandlung mit Azathioprin oder Cyclophosphamid;
4b	Studienorganisation und Ort der Studiendurchführung	7 Zentren in USA, 4 Zentren in Kanada
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Prüfsubstanzen</p> <ul style="list-style-type: none"> • Beta-Interferon-1b (IFNB, 50 µg, s. c.); • Beta-Interferon-1b (IFNB, 250 µg, s. c.); • Placebo

Item ^a	Charakteristikum	Studieninformation
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primäre Zielkriterien</p> <p>Unterschiede in der Schubrate und Anteil an Patienten, die schubfrei bleiben.</p> <p>Sekundäre Zielkriterien</p> <ul style="list-style-type: none"> • Zeit bis zum ersten Schub; • Dauer und Schwere des Schubs (definiert gemäß der quantitativen Veränderung im NRS-Wert: 0 bis 7 = mild; 8 bis 14 = mäßig; > 15 schwer); • Veränderung in EDSS und NRS-Werten zu Baseline; • Quantitative Messung der Krankheitslast durch jährliches MRT; • Messung der Krankheitsaktivität durch MRT (Substudie mit regelmäßigen Scans);
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Angaben*
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Keine Angaben*
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<p>Nach 2 Jahren wurde anhand der ersten 338 eingeschlossenen Patienten eine Intent-To-Treat Analyse durchgeführt. Allen Patienten, die in der Studie waren, wurde eine weitergehende doppelblinde Behandlung angeboten, die die Gesamtbehandlungszeit für einige Patienten auf 5 ½ Jahre ausweitete. Die Sicherheits- und Wirksamkeitsergebnisse für das erste Jahr dieser Extensionsstudie wurden ebenfalls in die 3-Jahresdaten eingeschlossen und diskutiert.</p> <p>Kriterien für einen vorzeitigen Ausschluss:</p> <ul style="list-style-type: none"> • Stetige Progression der Behinderung für 6 Monate; • Behandlung mit mehr als 3 Durchgängen von ACTH oder Prednison (alle 28 Tage oder weniger) während einer 1-Jahresperiode; • Nichteinnahme von geplanten Dosen des Studienmedikaments in mehr als 2 aufeinanderfolgenden Wochen; • Gemäßigte oder schwere Medikamententoxizität, die bei nochmaliger Untersuchung fort dauerte;
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Keine Angaben*
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Keine Angaben*
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur	Keine Angaben*

Item ^a	Charakteristikum	Studieninformation
	Zuteilung gewährleistet war	
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Keine Angaben*
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) Die Patienten waren verblindet. b) Jegliches Studienpersonal am Studienzentrum war verblindet. c) Zwei Ärzte an jedem Zentrum wurden festgelegt: ein Neurologe, der die Nebeneffekte der Prüfsubstanzen nicht kannte, um die periodischen Untersuchungen durchzuführen; und ein anderer behandelnder Neurologe, der die Nebeneffekte und die Reaktionen an der Einstichstelle kannte, die Laborauffälligkeiten nach Toxizität beurteilte und für die Gesamtversorgung verantwortlich war.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Identische Fläschchen mit lyophilisiertem IFNB (entweder 50 µg oder 250 µg) und Placebo wurden vorbereitet. Jedes IFNB-Fläschchen enthielt 15 mg USP humanes Albumin und 15 mg Dextrose zusätzlich zum IFNB-Inhalt und jedes Placebo-Fläschchen enthielt nur die entsprechende Menge an Albumin und Dextrose.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Wie im Studienprotokoll spezifiziert, wurden die Daten der beiden Studien kombiniert und den Analysen wurde der Intent-To-Treat Datensatz zugrunde gelegt. Die Analysen für die Extensionsstudien deckten die gesamten 3 Jahre an verfügbaren Daten ab. Für stetige Variablen wurden die Unterschiede in den Behandlungsgruppen unter Verwendung einer Varianz-Analyse (ANOVA), basierend auf gereihten Daten, analysiert. Dieses Modell ist für die folgenden Effekte ausgewiesen: Behandlungsgruppe, Studienzentrum und Behandlungsgruppe pro Studienzentrum. Diskrete Variablen, wie Geschlecht und Rasse, wurden auf allgemeine Zuordnung getestet, unter Verwendung eines Cochran-Mantel-Haenszel (CMH) Chi-Quadrat-Tests, stratifiziert nach Zentrum. Gleichmaßen wurden diskrete Variablen mit einer ordinalen Skalierung, wie Schubanzahl pro Patient, auf Unterschiede in der Behandlungsgruppe mit Hilfe einer CMH ANOVA Statistik getestet. Diese Statistik entspricht einer Zentrums-angepassten ANOVA. Ein zweiseitiger Fisher's Exact-Test wurde für die diskreten Variablen, die schwach waren, verwendet. In der Darstellung der Gruppen-Schubraten wurde das 95%-ige Konfidenzintervall unter Verwendung einer Poisson-Verteilung berechnet, basierend auf der Anzahl der beobachteten Schübe in jeder Behandlungsgruppe.

Item ^a	Charakteristikum	Studieninformation
		Überlebenskurven wurden mit der Sterbetafel-Methode, für die Dauer bis zum Ausbruch des ersten und zweiten Schubes, berechnet. Die Daten wurden ab dem Zeitpunkt eines Zurückziehens des Einverständnisses zensiert. Die Log-Rank Statistik wurde verwendet, um die Vergleichbarkeit der Überlebenskurven für jeden Behandlungsarm zu testen.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Sowohl die 2-Jahres, als auch die 3-Jahres-Daten wurden ausgewertet.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<p>Beta-Interferon-1b (50 µg, s. c., alle zwei Tage): a) 125 Patienten; b) 125 Patienten; c) 125 Patienten;</p> <p>Beta-Interferon-1b (250 µg, s. c., alle zwei Tage): a) 124 Patienten; b) 124 Patienten; c) 124 Patienten;</p> <p>Placebo: a) 123 Patienten; b) 123 Patienten; c) 123 Patienten;</p>
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Beta-Interferon-1b (50 µg, s. c., alle zwei Tage): Abbrüche insgesamt nach 2 Jahren: n = 18 1 wegen mangelnder Wirksamkeit 5 wegen unerwünschten Ereignissen</p> <p>Beta-Interferon-1b (250 µg, s. c., alle zwei Tage): Abbrüche insgesamt nach 2 Jahren: n = 24 1 wegen mangelnder Wirksamkeit 10 wegen unerwünschten Ereignissen</p> <p>Placebo: Abbrüche insgesamt nach 2 Jahren: n = 23 10 wegen mangelnder Wirksamkeit 1 wegen einem unerwünschten Ereignis</p> <p>Weitere Gründe für einen Abbruch in den 3 Behandlungsgruppen:</p> <ul style="list-style-type: none"> • Ereignisse, die eine Entblindung notwendig machten (n = 9) • Übermäßiger Gebrauch von Steroiden (n = 14) • wesentliche Non-Compliance (n = 2) • Schub an Baseline (n = 1) • emotionale Instabilität (n = 1) • persönliche Gründe (Ablehnung von Injektionen;

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

Item ^a	Charakteristikum	Studieninformation
		Umzug; der Wunsch, schwanger zu werden)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Patienten wurden von Juni 1988 bis Mai 1990 eingeschlossen. Die Studie wurde im Jahr 1995 beendet.
14b	Informationen, warum die Studie endete oder beendet wurde	Nach 2 Jahren wurde eine Intent-to-treat Analyse durchgeführt und die Studie wurde auf eine Gesamtzeit von 5 ½ Jahren verlängert. Die Studie wurde dann wie geplant abgeschlossen.
a: nach CONSORT 2010.		

* Keine Angaben hierzu in der Publikation.

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

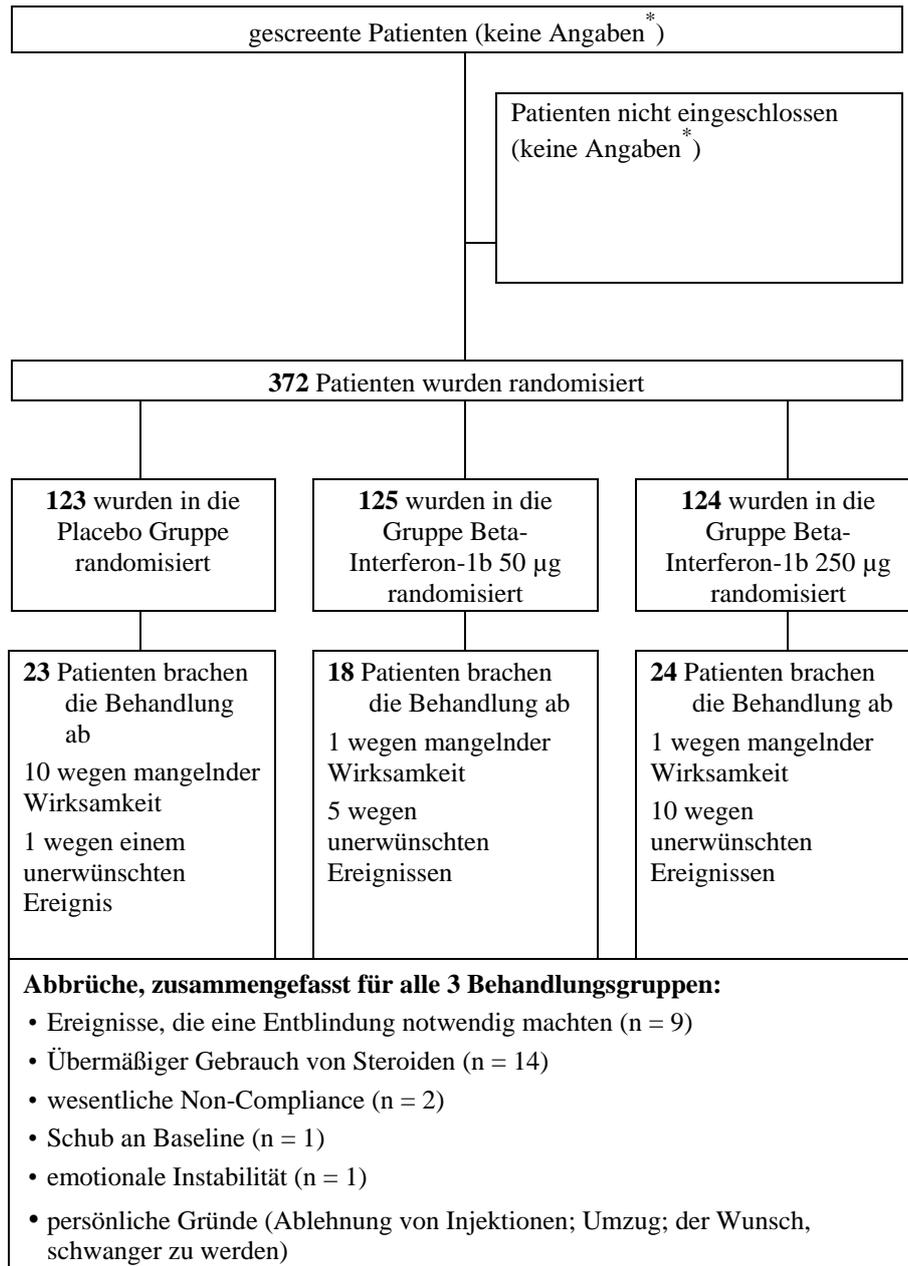


Tabelle 4-94 (Anhang): Studiendesign und -methodik für Studie MSCRG

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	<p>In dieser Studie wird untersucht, ob Beta-Interferon-1a die progressive, irreversible, neurologische Behinderung einer schubförmigen multiplen Sklerose verlangsamen könnte.</p> <p>Primärer Zielparameter:</p> <ul style="list-style-type: none"> • Dauer bis zur Verschlechterung der Behinderung, definiert als Verschlechterung zur Baseline um mindestens 1,0 Punkt in der EDSS-Skala, die Verschlechterung muss mindestens 6 Monate andauern; <p>Weitere Zielparameter:</p> <ul style="list-style-type: none"> • Anzahl und Volumen diskreter Gadolinium-angereicherter Läsionen im MRT; • Volumen von T2-Läsionen im MRT;
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	doppelblind, Placebo-kontrolliert und randomisiert; Efron's „Biased Coin“-Methode wurde für die Randomisierung benutzt.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Angaben *
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <ul style="list-style-type: none"> • Erwachsene (18 – 55 Jahre) mit eindeutig diagnostizierter multipler Sklerose seit mindestens einem Jahr; • mindestens zwei Schübe in den letzten drei Jahren; • keine Schübe innerhalb 2 Monate vor Studieneintritt • EDSS zur Baseline: 1,0 - 3,5; <p>Ausschlusskriterien:</p> <ul style="list-style-type: none"> • Vorherige immunsuppressive oder Interferon-Therapie oder Behandlung mit einem adrenokortikotropem Hormon oder Corticosteroiden innerhalb von 2 Monaten vor Studieneintritt; • Schwangerschaft oder Stillzeit; • Widerwille, Empfängnisverhütung einzuhalten; • Vorliegen einer chronisch-progressiven multiplen Sklerose; • Jegliche Erkrankung, die die Organfunktionen beeinträchtigt (außer MS);
4b	Studienorganisation und Ort der Studiendurchführung	<p>4 Zentren in USA</p> <p>Die Studienorganisation wurde von der „Multiple Sclerosis Collaborative Research Group“ (MSCRG) übernommen.</p>

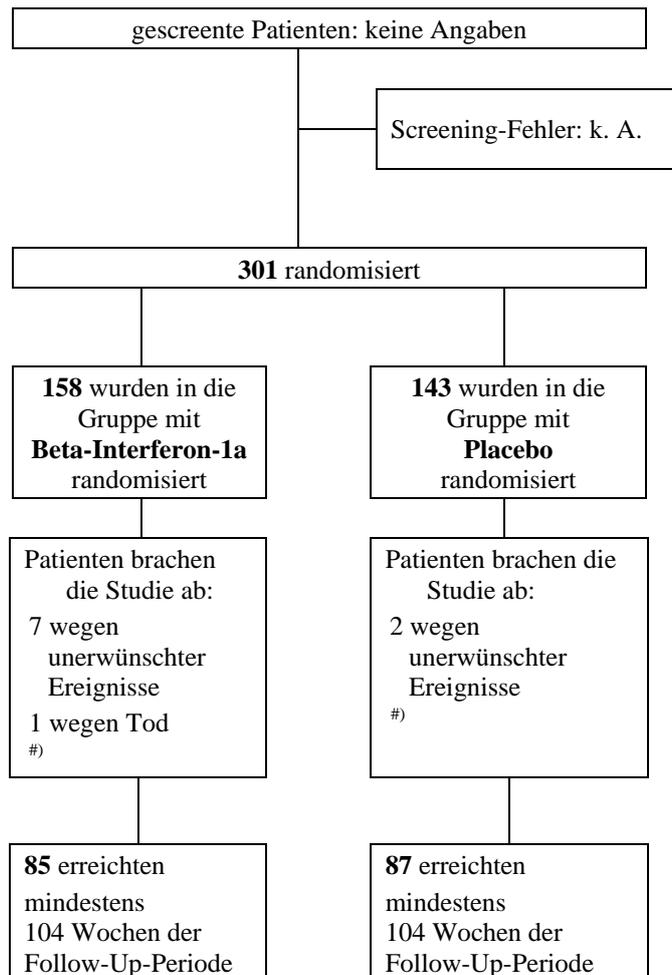
Item ^a	Charakteristikum	Studieninformation
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Prüfsubstanz:</p> <ul style="list-style-type: none"> • Beta-Interferon-1a (30 µg, i. m., einmal wöchentlich für bis zu 104 Wochen); <p>Die Injektionen wurden durch Studienschwestern oder durch lokale Mediziner unter Aufsicht des Studienpersonals durchgeführt.</p> <p>Kontrollsubstanz:</p> <ul style="list-style-type: none"> • Placebo
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Zielparameter:</p> <ul style="list-style-type: none"> • Dauer bis zur Verschlechterung der Behinderung, definiert als Verschlechterung zur Baseline um mindestens 1,0 Punkt in der EDSS-Skala; die Verschlechterung muss mindestens 6 Monate bestehen; <p>Eine Verschlechterung des EDSS konnte an einer geplanten oder ungeplanten Visite beginnen, aber sie musste über mindestens zwei geplante Visiten im Abstand von 6 Monaten fortauern, um vorübergehende Fluktuationen im klinischen Status auszuschließen. Die untersuchenden Ärzte mussten sich vor der Studie Trainingseinheiten unterziehen, um das Punktevergabeverfahren zu standardisieren.</p> <p>Weitere Zielparameter:</p> <ul style="list-style-type: none"> • Anzahl und Volumen diskreter Gadolinium-angereicherter Läsionen im MRT; • Volumen von T2-Läsionen im MRT; <p>Die Magnetresonanztomographie wurde jährlich unter Verwendung eines standardisierten Protokolls durchgeführt.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Angaben *
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Die Berechnung der Fallzahl basierte auf einer Kaplan-Meier-Analyse und einem „Intent-to-Treat“-Design. Es wurde angenommen, dass sich 50 % der Placebo-Empfänger und 33 % der Beta-Interferon-1a-Empfänger um mindestens 1,0 EDSS-Punkt innerhalb von 104 Wochen verschlechtern. Die erwartete Placebo-Progressionsrate orientierte sich am Median der Zeit für eine Progression (104 Wochen) im Placebo-Arm einer anderen klinischen Studie. Die Studie wurde mit einer statistischen Power von 80 % angelegt, um einen Gruppenunterschied mit einem α-Level von 0,05 nachzuweisen. Schätzungen zur Fallzahl gingen auch davon aus, dass 25 % der Beta-Interferon-1a-Empfänger die Behandlung vorzeitig abbrechen würden, aber immer noch zur Verlaufskontrolle erscheinen würden, während 10 % der Patienten keine Verlaufskontrolle nutzen würden.</p>

Item ^a	Charakteristikum	Studieninformation
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es wurden keine Zwischenanalysen durchgeführt. Die Entscheidung, die Studie früh zu beenden, wurde ohne die Kenntnis irgendwelcher Zwischenergebnisse hinsichtlich der Wirksamkeit getroffen. Diese Entscheidung wurde durch ein „Safety and Monitoring“-Komitee der „National Institutes of Health“ (NIH) begutachtet und genehmigt.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Efron's „Biased Coin“-Methode wurde für die Randomisierung benutzt.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Es erfolgte eine Stratifizierung nach Zentren.
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	Die Patienten wurden zufällig der Placebo- oder Beta-Interferon-1a-Gruppe zugeteilt. Keine sonstigen Angaben *
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Keine Angaben *
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) Die Patienten waren verblindet. b) Die Studienschwestern und die lokalen Mediziner, die die Injektionen verabreichten, waren verblindet. c) Die behandelnden und die unabhängigen untersuchende Ärzte waren verblindet. Die Patienten diskutierten keine medizinischen Fragen mit den untersuchenden Ärzten.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Keine Angaben *
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Für alle statistischen Analysen wurden zweiseitige Tests verwendet. Die Dauer bis zur beständigen Progression und bis zum ersten Schub wurden mittels der Kaplan-Meier-Methode analysiert, mit einer durch einen Log-Rang-Test ermittelten Signifikanz und stratifiziert nach den Zentren. Da die Studie auf einer Intent-to-Treat-Basis entworfen wurde, wurden die Patienten, die die Behandlung abbrachen, wann immer möglich, bis zum Ende der Studie nachbeobachtet und wurden in die Analysen mit einbezogen. Mehrere Analysen wurden mit dem Teil der Patienten durchgeführt, die früh genug hinzugekommen waren, um die 104 Wochen der Verlaufskontrolle am Ende der Studie abzuschließen.

Item ^a	Charakteristikum	Studieninformation
		<p>Dies beinhaltet den Anteil des Progressionsfortlaufs (Chi-Quadrat-Test), die individuelle Veränderung im EDSS (Mann-Whitney-Test) und die Anzahl der Schübe pro Patient (Mann-Whitney-Test). Jährliche Schubraten wurden auf dem üblichen Weg berechnet, indem die Gesamtzahl der Schübe während der ersten 104 Wochen durch die gesamten Personenjahre der Exposition geteilt wurde.</p> <p>Es wurden die üblichen Methoden zur Bestimmung der Signifikanzraten (z.B. der Wahrscheinlichkeitsquotiententest) verwendet. Jährliche Schubraten wurden ebenfalls für alle durchgängig an der Studie beteiligten Patienten auf die gleiche Weise wie oben berechnet.</p> <p>MRT-Analysen enthielten Vergleiche zu jedem Zeitpunkt über die Anzahl und dem Volumen Gadolinium-angereicherter Läsionen (Mann-Whitney-Test), dem Anteil angereicherter Scans (Chi-Quadrat-Test) und dem Prozentsatz der Veränderung im T2-Volumen zur Baseline (Mann-Whitney-Test).</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Keine Angaben*
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<p>Beta-Interferon-1a (30 µg, i. m., einmal wöchentlich):</p> <p>a) 158 Patienten; b) 158 Patienten; c) 85 Patienten;</p> <p>Placebo: a) 143 Patienten; b) 143 Patienten; c) 87 Patienten;</p>
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Beta-Interferon-1a (30 µg, i. m., einmal wöchentlich):</p> <p>Abbrüche insgesamt: n = k. A. 7 wegen unerwünschter Ereignisse 1 wegen Tod</p> <p>Placebo: Abbrüche insgesamt: n = k. A. 2 wegen unerwünschter Ereignisse</p> <p>Insgesamt waren 5 Patienten nicht mehr für die Verlaufskontrolle verfügbar.</p>
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Studie wurde im November 1990 initiiert und im Jahr 1995 beendet.

Item ^a	Charakteristikum	Studieninformation
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie konnte etwa 1 Jahr früher enden als geplant, ohne statistische Power einzubüßen, da die Abbrecherquote der Patienten mit etwas weniger als 3 % deutlich unter der ursprünglich geschätzten Abbrecherquote von 10 % lag.
a: nach CONSORT 2010.		

* Keine Angaben hierzu in der Publikation.



#) Insgesamt waren 5 Patienten nicht mehr für die Verlaufskontrolle verfügbar.

Tabelle 4-95 (Anhang): Studiendesign und -methodik für Studie Johnson et al., 1995

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	<p>Grundhypothese: Bestätigung der Schlussfolgerungen der Pilotstudie von Bornstein et al. (1987).</p> <p>Primärziel: Vergleich der mittleren Schubhäufigkeiten bei Patienten mit RRMS nach zweijähriger Behandlung mit Glatirameracetat (20 mg; s.c. tgl.) oder passendem Placebo.</p> <p>Sekundärziele: Vergleich der beiden Behandlungsgruppen von Randomisierung bis Studienende: - Anteil der Patienten ohne Schübe; - Zeit bis zum Auftreten des ersten Schubs nach Behandlungsbeginn; - Patientenanteil mit anhaltender Progression; - Mittlere Veränderungen in EDSS und im Ambulation-Index.</p>
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Prospektiv, doppelblind, randomisiert, Placebo-kontrolliert, parallel. Randomisierungsverhältnis 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Angaben*
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <ol style="list-style-type: none"> 5. Klinisch eindeutige oder durch Laborparameter abgesicherte MS-Diagnose 6. Männer und Frauen (18-45 Jahre) 7. EDSS-Score 0 - 5,0 8. mindestens zwei eindeutig diagnostizierte und dokumentierte Schübe innerhalb der letzten zwei Jahre vor Studienbeginn 9. der erste Schub mind. ein Jahr vor Einschluss 10. 30 Tage vor Randomisierung neurologisch stabil und ohne Kortikosteroidbehandlung <p>Ausschlusskriterien:</p> <ol style="list-style-type: none"> 1. vorherige Behandlung mit Glatirameracetat 2. vorherige immunsupprimierende zytotoxische Chemotherapie (z.B. Azathioprin, Cyclophosphamid oder Ciclosporin) 3. vorherige Bestrahlung von Lymphorganen 4. Schwangere oder stillende Frauen 5. Insulin-abhängige Diabetes Mellitus

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

Item ^a	Charakteristikum	Studieninformation
		<p>6. Positive HIV oder HTLV-I Serologie</p> <p>7. Hinweise auf Lyme-Krankheit (Borreliose)</p> <p>8. Bedarf einer Behandlung mit ASS oder NSAIDs im Verlauf der Studie</p> <p>9. keine Anwendung adäquater Kontrazeptionsmethoden (Frauen)</p>
4b	Studienorganisation und Ort der Studiendurchführung	USA, 11 Zentren (Universitätskliniken mit MS-Zentren) unter Leitung der Universität von Maryland.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Prüfsubstanz: Glatirameracetat (Copolymer 1/Copaxone) Lyophilisat und steriles Wasser als Lösungsmittel 20 mg s.c. 1x tgl.</p> <p>Kontrollsubstanz: Placebo Lyophilisat und steriles Wasser als Lösungsmittel</p> <p>Die Patienten erhielten je einen Vorrat für einen Monat und zu Beginn der Studie eine detaillierte Einführung in Rekonstitution und s.c. Eigenapplikation.</p>
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärziel: Anzahl Schübe nach zweijähriger Behandlung mit Glatirameracetat (20 mg; s.c. tgl.) oder passendem Placebo.</p> <p>Schub: Auftreten oder Wiederauftreten einer oder mehrerer neurologischer Auffälligkeiten für mindestens 48 Stunden und zuvor eine mindestens 30-tägige Phase mit relativ stabilem oder sich verbesserndem neurologischen Status. Eine Bestätigung des Schubs erfolgte nur bei gleichzeitiger objektiver Verschlechterung des Befindens (z.B. EDSS).</p> <p>Sekundärziele: Vergleich der beiden Behandlungsgruppen von Randomisierung bis Studienende:</p> <ul style="list-style-type: none"> - Anteil der Patienten ohne Schübe; - Zeit bis zum Auftreten des ersten Schubs nach Behandlungsbeginn; - Patientenanteil mit anhaltender Progression; - Mittlere Veränderungen in EDSS und im Ambulation-Index. <p>Vierteljährliche neurologische Untersuchung (verblindete Neurologen und Studienschwestern), ständige Erfassung von Schüben</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Angaben*
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Keine Angaben*
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Vierteljährliche Sitzungen eines verblindeten, unabhängigen Safety Monitoring Komitees (zwei Neurologen, ein Pharmakologe, ein Statistiker und ein

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

Item ^a	Charakteristikum	Studieninformation
		Patientenvertreter [National MS Society]) zur Beurteilung der Sicherheitsaspekte.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Keine Angaben*
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Keine Angaben*
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	Zentralisiertes Randomisierungsschema
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Patienten wurden in den Zentren von den jeweiligen Prüfarzten in die Studie aufgenommen und mittels einer zentralisierten Randomisierungsliste einem Behandlungsarm zugeordnet.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) Die Patienten waren verblindet. b) Das für die Ausgabe der Medikation, Erfassung von Begleittherapien und die Entnahme von Blut- und Urinproben zuständige Zentrumspersonal war verblindet. c) Für die Erfassung der neurologischen Befunde kam vierteljährlich ein „Zwei-Neurologen-Protokoll“ zur Anwendung: ein Neurologe erfasste den objektiven neurologischen Zustand ohne Diskussion von Nebenwirkungen oder Symptomen, ein zweiter Neurologe erfasste die Symptome und AEs und beurteilte die Notwendigkeit zur Steroidbehandlung bei Schüben; beide Neurologen waren verblindet.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Keine Angaben*
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Primärziel: Kovarianzanalyse (ANCOVA) mit Tests für Wechselwirkungen zwischen Studienmedikament und Zentren Sekundärziele: - Anzahl Patienten ohne Schübe: Logistische Regressionsanalyse - Zeit bis zum Auftreten des ersten Schubs nach Behandlungsbeginn: Weibull-Regression - Patientenanteil mit anhaltender Progression: logistische Regression - Veränderungen in EDSS und Ambulation-Index: Kovarianzanalyse mit Messwiederholungen (repeated-measured ANCOVA)

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

Item ^a	Charakteristikum	Studieninformation
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Primärziel: Einteilung in vordefinierte Gruppen: Geschlecht, Erkrankungszeit, Schubhäufigkeit in den zwei Jahren vor Einschluss und EDSS bei Baseline (0-2, 2-4, >4).
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Glatirameracetat: a) 125 Patienten b) 125 Patienten erhielten Glatirameracetat, 106 davon „per protocol“ mindestens 24 Monate lang c) 125 Patienten („intent-to-treat“-Kohorte) Placebo: a) 126 Patienten b) 126 Patienten erhielten Placebo, 109 davon „per protocol“ mindestens 24 Monate lang c) 126 Patienten („intent-to-treat“-Kohorte)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Glatirameracetat: Abbrüche insgesamt: n = 19 3 Patientinnen wegen Schwangerschaft 1 Patient: MS-Progression 2 Patienten: auftreten eines SAE 3 Patienten: Rücknahme der Einwilligung wegen AE Placebo: Abbrüche insgesamt: n = 17 2 Patienten: fehlende Compliance zum Prüfplan 1 Patient: Rücknahme der Einwilligung wegen AE Ansonsten keine Informationen vorhanden.
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Patienten wurden von Oktober 1991 bis Mai 1992 eingeschlossen.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde planmäßig beendet
a: nach CONSORT 2010.		

* Keine Angaben hierzu in der Publikation.

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

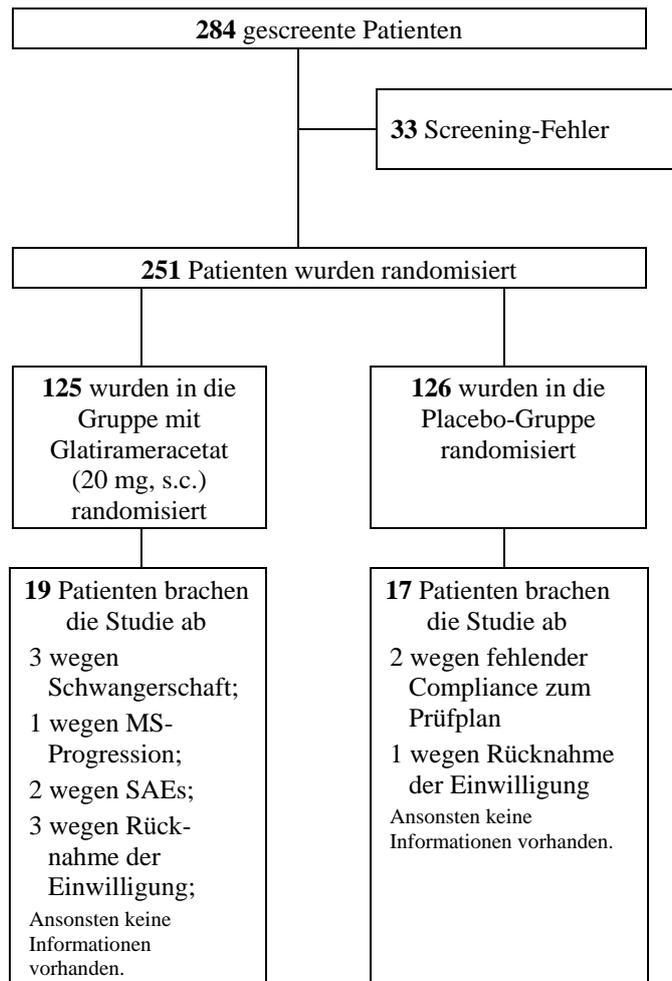


Tabelle 4-96 (Anhang): Studiendesign und -methodik für Studie REGARD

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	<p>Dies ist eine vergleichende Studie zwischen Beta-Interferon-1a und Glatirameracetat bei Patienten mit RRMS.</p> <p>Primärer Zielparameter:</p> <ul style="list-style-type: none"> • Dauer bis zum ersten Schub über 96 Wochen, als „Hazard Ratio“ (HR) ausgedrückt; <p>Sekundäre Zielparameter:</p> <ul style="list-style-type: none"> • Mittlere Anzahl neuer aktiver T2-Läsionen; • Mittlere Anzahl Gadolinium-angereicherter Läsionen; • Veränderungen im Volumen der Gadolinium-angereicherten und der T2-Läsionen;
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Multi-zentrische, randomisierte, parallele und offene Studie; Randomisierung 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Angaben *
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <ul style="list-style-type: none"> • Erwachsene (18 – 60 Jahre) mit RRMS, die Beta-Interferon- und Glatirameracetat-naive waren; • mindestens ein Schub im letzten Jahr; • klinisch stabil oder sich neurologisch verbessernd innerhalb der 4 Wochen vor Randomisierung; • EDSS zwischen 0 und 5,5; <p>Ausschlusskriterien:</p> <ul style="list-style-type: none"> • Schwangerschaft oder Stillzeit; • Progressive MS; • Behandlung mit Steroiden (oral oder systemisch) oder mit adrenokortikotropem Hormon innerhalb der letzten 4 Wochen; • frühere Behandlung mit Beta-Interferon, Glatirameracetat oder Cladribin; • Komplette Lymphbestrahlung; • Plasma-Austausch innerhalb der letzten 3 Monate; • Gebrauch von intravenösem Gamma-Globulin innerhalb der letzten 6 Monate; • Cytokin- oder Anticytokin-Therapie innerhalb der letzten 3 Monate; • Gebrauch eines Immunsuppressivums innerhalb der letzten 12 Monate;
4b	Studienorganisation und Ort der Studiendurchführung	81 Zentren in 14 Ländern (Argentinien, Brasilien, Deutschland, Frankreich,

Item ^a	Charakteristikum	Studieninformation
		Irland, Italien, Kanada, Niederlande, Österreich, Russland, Schweiz, Spanien, USA und Vereinigtes Königreich); Das Studienprotokoll wurde durch den Sponsor und durch einen Lenkungsausschuss der Prüfer entwickelt.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Prüfsubstanzen: • Beta-Interferon-1a (44 µg, s. c., dreimal pro Woche); • Glatirameracetat (20 mg, s. c., täglich); Die Patienten wurden zufällig zugeteilt, um die jeweilige Prüfsubstanz vom Zentrumspersonal zu erhalten.
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	Primärer Zielparameter: • Dauer bis zum ersten Schub über 96 Wochen, als „Hazard Ratio“ (HR) ausgedrückt; Sekundäre Zielparameter: • Mittlere Anzahl neuer aktiver T2-Läsionen (definiert als neue oder vergrößerte T2-Läsion pro Patient pro Scan über 96 Wochen); • Mittlere Anzahl Gadolinium-angereicherter Läsionen pro Patient pro Scan; • Veränderungen im Volumen der Gadolinium-angereicherten und der T2-Läsionen; Die Patienten erhielten am Behandlungstag 1 ihre erste Studienmedikamentengabe. Klinische Beurteilungen im Rahmen der Verlaufskontrolle wurden 4, 12 und 24 Wochen und alle 24 Wochen danach bis zu 96 Wochen durchgeführt. Eine Untergruppe von 460 Patienten erhielt fortlaufende Gehirn-MRT-Scans an den Wochen 24, 48, 72 und 96. MRT-Bewertungen wurden zentral und verblindet am „Image Analysis Center“ der Freien Universität in Amsterdam durchgeführt.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Angaben*
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Um mit einer 80 %-igen Power einen signifikanten Unterschied in der Dauer bis zum ersten Schub zu ermitteln, wurden 736 Patienten in jedem Arm benötigt. Diese Berechnung basierte auf einem zweiseitigen Log-Rang-Test, unter der Annahme eines Fehlers der 1. Art von 5 % und einer erwarteten HR von 0,77, was eine 30 % längere Dauer bis zum ersten Schub ergibt, wenn der Median der Dauer bis zum ersten Schub vermutlich 600 Tage in der Beta-Interferon-1a- und 462 Tage in der Glatirameracetat-Gruppe beträgt.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Eine verblindete Zwischenanalyse der Gesamtzahl der Schübe und der Dauer bis zum ersten Schub wurde durchgeführt, nachdem alle Patienten ihre

Item ^a	Charakteristikum	Studieninformation
		jeweilige Behandlung für 72 Wochen erhalten hatten. MRT-Scans wurden bei 60 % der Patienten beurteilt, damit eine 85 %-ige statistische Power zur Verfügung steht, einen 40 %-igen Unterschied in der mittleren Anzahl der T2-aktiven Läsionen pro Patient pro Scan während der 96 Behandlungswochen zu ermitteln. Diese Berechnung wurde mit einem zweiseitigen Wilcoxon-Rangsummentest mit einem 5 %-igen Fehler der 1. Art durchgeführt.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die zufällige Zuteilung wurde mittels Computer generiert.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierungsliste wurde nach Zentren 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	Keine Angaben *
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Behandlungen wurden nach einer computer-generierten Randomisierungsliste zugeteilt.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) Die Patienten waren nicht verblindet. b) Die Patienten erhielten die Medikation vom behandelnden Studienpersonal. c) Die beurteilenden Ärzte waren verblindet. Die Ärzte, die die Patienten in regulären Intervallen oder zum Zeitpunkt eines möglichen Schubes beurteilten, waren gegenüber der Behandlung verblindet und kommunizierten mit dem Patienten nur insoweit, um EDSS, die Kurtzke Funktionsskala und die Schub-Burteilungen abschließen zu können. Patienten wurden gebeten, ihre Behandlung nicht mit dem beurteilenden Arzt zu besprechen und ihre Injektionsstellen zu bedecken, so dass der Arzt nicht vermuten konnte, welche Behandlung sie erhalten hatten. Keine weiteren Angaben *
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht nötig, da die Patienten nicht verblindet waren.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die primäre Analyse wurde mit der Intent-To-Treat-Population durchgeführt. Patienten, die 96 Wochen der Behandlung und der Beurteilungen ohne größere Protokollabweichungen erfüllten, wurden in die Per-

Item ^a	Charakteristikum	Studieninformation
		<p>Protokoll-Population mit einbezogen.</p> <p>Stetige Variablen zur Baseline wurden mit einem zweiseitigen Varianzanalysemodell mit gereihten Daten beurteilt, das Behandlungs- und Zentrumseffekte berücksichtigte. Diskrete Variablen wurden mit einem allgemeinen Cochran-Mantel-Haenszel Assoziations- oder Reihen-Mittelwert-Test beurteilt, die auf das Zentrum abgestimmt waren.</p> <p>Der primäre Endpunkt nach 96 Wochen wurde mit dem zu Cox proportionalem Hazard-Modell analysiert, mit Behandlungs- und Zentrumseffekten. Die HR (95 % KI) der Beta-Interferon-1a Gruppe gegen die Glatirameracetat-Gruppe wurde vom zu Cox proportionalen Hazard Model mit der Glatirameracetat-Gruppe als Referenzgruppe berechnet, und die Mediane (oder verschiedene Perzentilen) der Dauer bis zum ersten Schub wurden aus Kaplan-Meier Überlebenskurven geschätzt. Die mittlere Anzahl der T2-aktiven Läsionen pro Patient pro Scan während der 96 Behandlungswochen wurde mit einer nicht-parametrischen Kovarianzanalyse (ANCOVA) nach gereihten Daten analysiert, mit Behandlungs- und Zentrumseffekten und abgestimmt auf Gadolinium-angereicherte Läsionszählungen zur Baseline. Die mittlere Zahl neuer T1 hypointenser Läsionen pro Patient pro Scan und alle stetigen MRT-Daten und die Veränderung im Gehirnvolumen über 2 Jahre wurden mit einer ähnlichen Annäherung analysiert. Der Anteil der Patienten mit keinen neuen T2-Läsionen und keinen Gadolinium-angereicherten Läsionen (oder andere binäre diskrete Daten) wurden mit einem logistischen Regressionsmodell mit Behandlungs- und Zentrumseffekten analysiert.</p> <p>Schubzählungen wurden mit einem Poisson-Regressionsmodell mit Behandlungs- und Zentrumseffekten analysiert (der Logarithmus der bei der Studie verbrachten Zeit wurde als Offset-Variable in dem Modell benutzt). Die jährliche Schubrate wurde für jede Behandlungsgruppe als Gesamtzahl der Schübe über alle Patienten geteilt durch die gesamte bei der Studie verbrachte Zeit über alle Patienten geschätzt. Der Anteil der Patienten, die am Ende der 96 Wochen schubfrei waren, der Anteil der Patienten mit einer bestätigten Progression nach 6 Monaten, der Anteil der Patienten mit keinen T2-aktiven Läsionen und der Anteil der Patienten mit keinen neuen T1-hypointensen Läsionen wurden mit einem logistischen Regressionsmodell analysiert, das Behandlungs- und Zentrumseffekte berücksichtigte.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Die MRT-Intent-To-Treat-Untergruppe umfasste 460 Patienten (230 Patienten in jedem Behandlungsarm).
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	

Item ^a	Charakteristikum	Studieninformation
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Beta-Interferon-1a (44 µg, s. c., dreimal pro Woche): a) 386 Patienten; b) 383 Patienten; c) 386 Patienten; Glatirameracetat (20 mg, s. c., täglich): a) 378 Patienten; b) 375 Patienten; c) 378 Patienten;
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Beta-Interferon-1a (44 µg, s. c., dreimal pro Woche): Abbrüche insgesamt: n = 85 23 wegen unerwünschter Ereignisse 8 wegen Schwangerschaft 2 wegen Protokollverletzungen 4 wegen Erkrankungs-Progression 28 aus anderen Gründen 3 erhielten kein Studienmedikament 17 ohne oder ohne komplette aktuelle Verlaufskontrolle Glatirameracetat (20 mg, s. c., täglich): Abbrüche insgesamt: n = 54 19 wegen unerwünschter Ereignisse 5 wegen Schwangerschaft 2 wegen Protokollverletzungen 7 wegen Erkrankungs-Progression 16 aus anderen Gründen 3 erhielten kein Studienmedikament 2 ohne komplette aktuelle Verlaufskontrolle
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Rekrutierungsphase der Studie begann im Februar 2004 und endete im Dezember 2004. Im September 2007 wurde die Studie beendet.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde wie geplant abgeschlossen.
a: nach CONSORT 2010.		

* Keine Angaben hierzu in der Publikation.

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

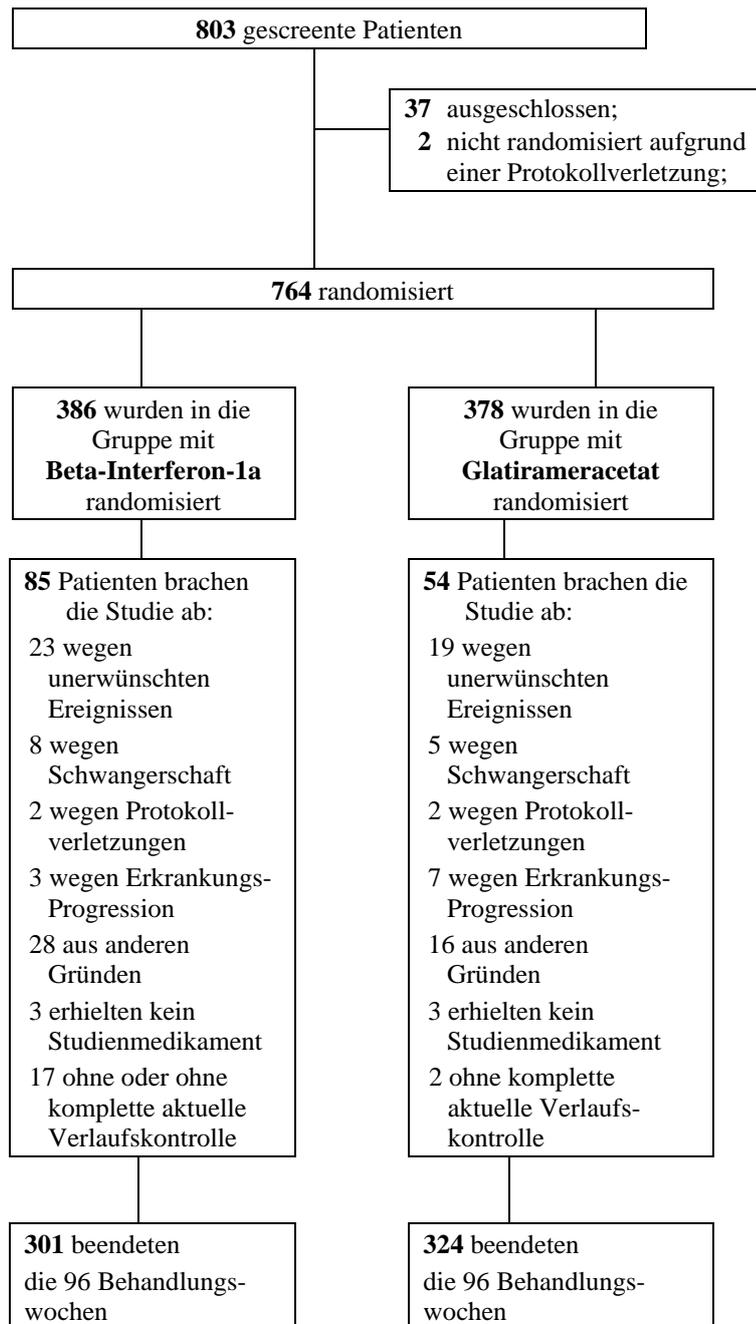


Tabelle 4-97 (Anhang): Studiendesign und -methodik für Studie BEYOND

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	Vergleich der Wirksamkeit, Sicherheit und Verträglichkeit von zwei Beta-Interferon-1b Dosierungen: <ul style="list-style-type: none"> ○ 250 µg (Standarddosis) ○ 500 µg mit Glatirameracetat bei der Behandlung von RRMS im Hinblick auf klinische Messungen und MRT-Messungen.
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Phase 3, prospektiv, multizentrisch, randomisiert, parallel Doppelblind für die Interferon-Dosierung, einfachblind für die Bewertung; Randomisierungsverhältnis: 2:2:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Auf Grund von zu niedrigen Ereignisraten wurde ca. 2 Jahre nach Behandlungsbeginn des ersten Patienten ein „filling of the triangle“-Design eingeführt. Damit wurden die Visiten am Prüfzentrum für die anfänglich eingeschlossenen Patienten solange fortgeführt, bis der letzte Patient den 2-Jahresendpunkt erreicht hatte.
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: <ul style="list-style-type: none"> - Erwachsene (18 – 55 Jahre) mit (früher) RRMS; - Behandlungs-naiv; - mind. 1 Schub im letzten Jahr; - EDSS 0 – 5,0; - bei gebärfähigen Frauen: negativer Schwangerschaftstest und adäquate Kontrazeption während der Studie; Ausschlusskriterien: <ul style="list-style-type: none"> - Anzeichen oder Symptome, welche eine andere Erkrankung als MS vermuten lassen; - progressive MS-Formen; - Herzerkrankung; - Vorbehandlung oder Teilnahme an früheren klinischen MS-Studien; - schwere Depression; - Alkohol- oder Drogenmissbrauch; - Suizidversuch oder Suizidgedanken; - schwere oder akute Fehlfunktion der Leber, der Nieren oder des Knochenmarks; - Gamma-Globinopathie;

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> - unkontrollierte Epilepsie; - Intoleranz, Kontraindikation oder Allergie gegenüber irgendeiner Studienmedikation; - MRT nicht durchführbar; - Verabreichung der Studienmedikation durch Patient oder Pflegekraft nicht möglich;
4b	Studienorganisation und Ort der Studiendurchführung	198 Zentren in 26 Ländern weltweit;
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Prüfsubstanz:</p> <ul style="list-style-type: none"> - 250 µg Beta-Interferon-1b, s. c.; alle 2 Tage; - 500 µg Beta-Interferon-1b, s. c., alle 2 Tage; - 20 mg Glatirameracetat, s. c., 1x tgl.; <p>Alle Patienten mussten einen Autoinjektor verwenden.</p> <p>zur Beta-Interferon-1b-Dosierung: Dosis-Titration zur Optimierung der Verträglichkeit: Erhöhung um 0,25 ml alle 2 Wochen bis zum Erreichen der Enddosis nach 6 Wochen (Ausdehnung auf 13 Wochen möglich).</p>
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>primärer Zielparameter: Schubrisiko (Definition <u>Schubrisiko</u>: neue oder rezidivierende neurologische Symptome mit einem Mindestabstand von 30 Tagen zum vorangehenden Ereignis und einer Dauer von mindestens 24 h).</p> <p>Sekundäre Zielparameter:</p> <ul style="list-style-type: none"> - EDSS-Progression; - Relative Veränderung im T1-hypointensen Läsionsvolumen vom Screening bis zum letzten Scan; <p>Die klinischen Zielparameter wurden vierteljährlich für die Dauer von 2 – 3,5 Jahre bestimmt.</p> <p>MRT-Messungen: wurden durchgeführt bei Baseline und anschließend jährlich. Es wurden T2-gewichtete und T1-gewichtete Bilder (mit und ohne DTPA) des gesamten Gehirns gemacht. T1-gewichtete und T2-gewichtete Läsionen wurden durch Auswerter identifiziert und das Läsionsvolumen wurde mit Hilfe einer halb-automatischen, lokalen Schwellenwert-Technik quantifiziert. Das normalisierte Gehirnvolumen zur Baseline wurde mit SIENAX („Structural Image Evaluation, Using Normalisation, of Atrophy: State“) beurteilt. Die prozentuale Änderung im longitudinalen Gehirnvolumen wurde mit SIENA („Structural Image Evaluation, Using Normalisation, of Atrophy: Rate“) beurteilt.</p>

Item ^a	Charakteristikum	Studieninformation
		Ein neurologisches Ereignis wurde nur als <u>Schub</u> erachtet, wenn es mit einem Anstieg in EDSS oder funktionalen Systemwerten assoziiert war und mit den beschriebenen Symptomen übereinstimmte. Zur Ermittlung der Schwere eines Schubs wurde ein Algorithmus entwickelt, welcher zwischen schweren und leichten Schüben unterschied. EDSS-Scores wurden erhoben bei neuen, wiederkehrenden oder sich verschlechternden neurologischen Symptomen sowie bei den geplanten Visiten.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Für die Berechnung der Power wurde ein Risiko-Quotient für wiederauftretende Schübe von 0,825 für 500 µg Beta-Interferon-1b und von 1,244 für Glatirameracetat relativ zu 250 µg Beta-Interferon-1b bestimmt. Wenn mindestens 840 Patienten zu jeder Beta-Interferon-1b-Dosierung und mindestens 420 Patienten zu Glatirameracetat zugeordnet werden, wird eine mindestens 80 %-ige Power für die Detektion von Unterschieden dieser Größenordnung zwischen den Gruppen gewährleistet. Die Berechnungen wurden mit der Software SAS (Version 9.1.3) durchgeführt.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Ein unabhängiges DSMB führte Zwischenanalysen zur Beurteilung des Nutzen-Risiko-Verhältnisses der Behandlungen durch.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	zentrale Gruppe zur Durchführung der Randomisierungen
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	SAS-basierte Blockrandomisierung mit regionaler Stratifizierung
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Keine Angaben*
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Keine Angaben*
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten,	Verblindung hinsichtlich der Zuordnung zu Beta-Interferon-1b oder zu Glatirameracetat: a) Die Patienten waren nicht verblindet.

Item ^a	Charakteristikum	Studieninformation
	verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	<p>b) Das behandelnde Studienpersonal war nicht verblindet.</p> <p>Verblindung hinsichtlich der Beta-Interferon-1b-Dosierung:</p> <p>a) Die Patienten waren verblindet.</p> <p>b) Das behandelnde Studienpersonal war verblindet.</p> <p>Generell:</p> <p>c) Die bewertenden Ärzte, die alle neurologischen Untersuchungen durchführten und die funktionalen System- und EDSS-Werte aufnahmen, waren im Bezug auf alle Behandlungsarme verblindet. Die bewertenden Ärzte waren nicht in die Behandlung des Patienten involviert und hatten keinen Zugang zu den Patientenakten oder früheren Untersuchungen.</p> <p>Die Patienten bedeckten ihre Injektionsstellen während der neurologischen Untersuchung und diskutierten keine unerwünschten Ereignisse mit dem bewertenden Arzt.</p>
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	<p><u>Beta-Interferon-1b-Dosierungen:</u> identisches Aussehen, Verpackung und Kennzeichnung der zwei Medikationsdosierungen;</p>
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Schub-Risiko:</p> <p>Andersen-Gill Model für die Dauer bis zum Wiederauftreten des Ereignisses (inkl. Korrektur für die Zwischenabhängigkeit multipler Ereignisse)</p> <p><u>Unterstützende Schub-basierte Ziele:</u></p> <ul style="list-style-type: none"> - <u>Dauer bis zum ersten Schub:</u> zu Cox proportionales Risiko-Modell; - <u>Anteil der für 2 Jahre schubfreien Patienten:</u> Kaplan-Meier-Schätzer - <u>Jährliche Schub-Rate:</u> Risiko-Quotient, abgeleitet von einer generalisierten linearen Poisson-Regression <p>MRT-Variablen</p> <p>(z.B. Entwicklung neuer T2-Läsionen, relative Änderung in T2-Läsionsvolumen, Änderung im Gehirnvolumen, Änderung im Volumen und Anzahl Gadolinium-anreichernder Läsionen): Beurteilung durch ein nichtparametrisches Regressionsmodell mit Baseline-Scan als eine Kovariate.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Zur Bewertung der Robustheit der Behandlungen in Bezug auf das Schub-Risiko wurden vordefinierte Subgruppenanalysen durchgeführt. Diese beinhalteten Stratifizierungen nach Altersmedian (Jahre), Geschlecht, Median der Krankheitsdauer (Jahre), geographische Region, Baseline EDSS Wert (≤ 2 oder > 2), Anzahl der Schübe im Jahr vor dem Einschluss (1, 2 oder mehr), Gadolinium-anreichernde Läsionen beim Screening, Median des T2-Läsionsvolumens beim Screening und diagnostische Kategorien (1, 2 oder mehr Schübe mit</p>

Item ^a	Charakteristikum	Studieninformation
		einer oder mehr klinischen Läsionen). <u>Häufigkeiten unerwünschter Ereignisse und schwerer unerwünschter Ereignisse</u> : Zusammenfassende Statistik (z.B. Anzahl und Mittelwert) und Analyse mit Fisher's Exact-Test.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	a) 2244 Patienten wurden randomisiert. - 899 zur IFN-β-1b 500 µg Gruppe - 897 zur IFN-β-1b 250 µg Gruppe - 448 zur Glatirameracetat 20mg Gruppe b) 2220 Patienten erhielten mindestens eine Dosis der geplanten Behandlung. c) mindestens 1884 Patienten, die die Studie vollständig beendet haben, wurden analysiert; 2096 Patienten wurden in die letzte verfügbare Scan-Analyse eingeschlossen; sonst keine weiteren Angaben* Beta-Interferon-1b (250 µg, s. c.; alle 2 Tage): a) 897 Patienten; b) 888 Patienten; c) Analyse per Protokoll, 784 Patienten haben die Studie vollständig abgeschlossen, sonst keine weiteren Angaben*; Beta-Interferon-1b (500 µg, s. c.; alle 2 Tage): a) 899 Patienten; b) 887 Patienten; c) Analyse per Protokoll, 726 Patienten haben die Studie vollständig abgeschlossen, sonst keine weiteren Angaben*; Glatirameracetat (20 mg, s. c., 1x tgl.): a) 448 Patienten; b) 445 Patienten; c) Analyse per Protokoll, 374 Patienten haben die Studie vollständig abgeschlossen, sonst keine weiteren Angaben*;
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Beta-Interferon-1a (250 µg, s. c.; alle 2 Tage): Abbrüche insgesamt: n = 104 10 wegen Behandlungswechsel 38 wegen Zurückziehens des Einverständnisses 3 wegen Protokollverletzungen 13 wegen unerwünschter Ereignisse 12 ohne aktuelle Verlaufskontrolle 1 wegen Schwangerschaft 27 aus anderen Gründen

Item ^a	Charakteristikum	Studieninformation
		<p>Beta-Interferon-1a (500 µg, s. c.; alle 2 Tage): Abbrüche insgesamt: n = 161 12 wegen Behandlungswechsel 53 wegen Zurückziehens des Einverständnisses 2 wegen Protokollverletzungen 20 wegen unerwünschter Ereignisse 3 wegen Tod 20 ohne aktuelle Verlaufskontrolle 3 wegen Schwangerschaft 48 aus anderen Gründen</p> <p>Glatirameracetat (20 mg, s. c., 1x tgl.): Abbrüche insgesamt: n = 71 5 wegen Behandlungswechsel 18 wegen Zurückziehens des Einverständnisses 2 wegen Protokollverletzungen 8 wegen unerwünschter Ereignisse 1 wegen Tod 12 ohne aktuelle Verlaufskontrolle 1 wegen Schwangerschaft 24 aus anderen Gründen</p>
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Screening: November 2003 – Juni 2005; Die Studie wurde im August 2007 beendet Eine Nachbeobachtung erfolgte mindestens für 2 Jahre.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde wie geplant abgeschlossen.
a: nach CONSORT 2010.		

* Keine Angaben hierzu in der Publikation.

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

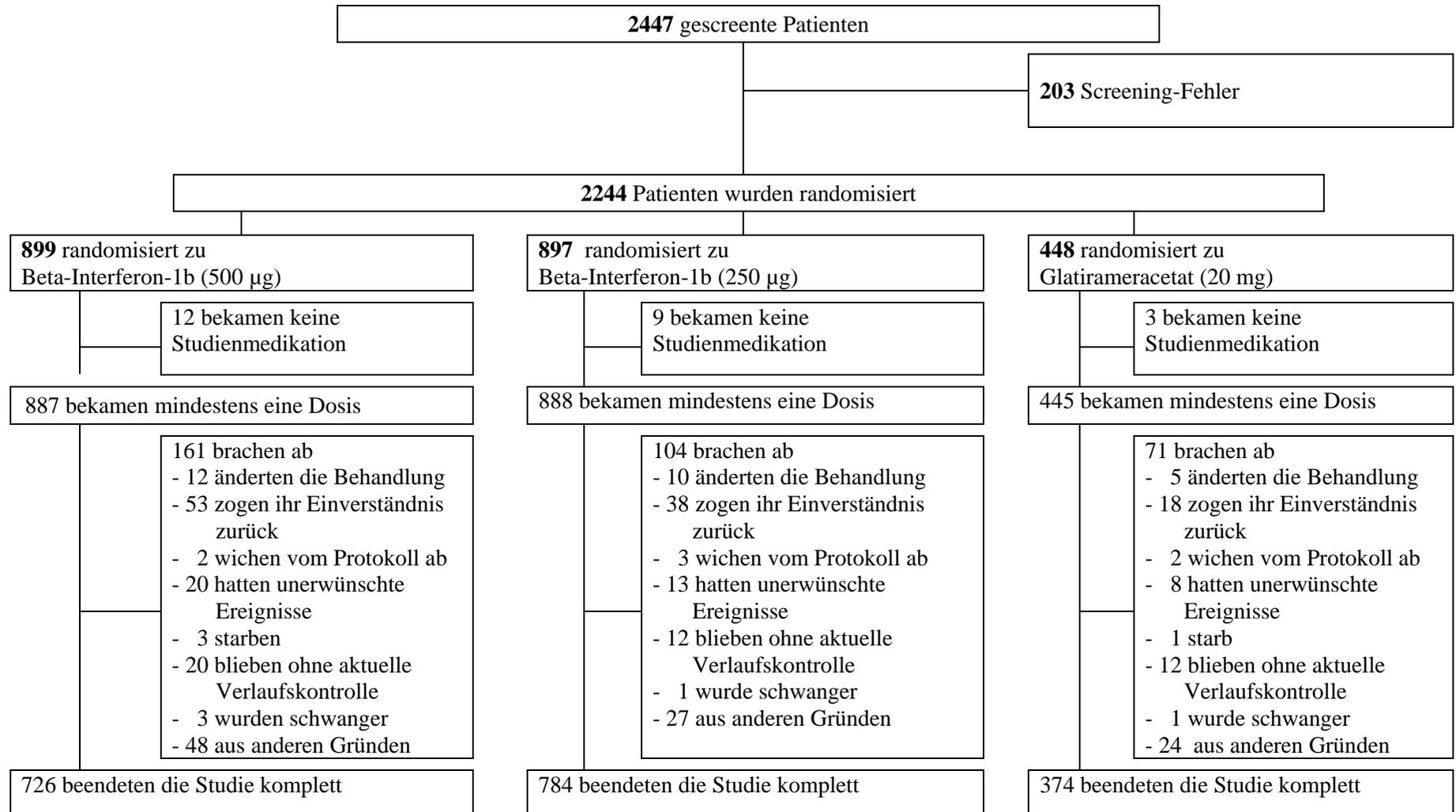


Tabelle 4-98 (Anhang): Studiendesign und -methodik für Studie EVIDENCE

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Vergleich der Wirksamkeit und Sicherheit von zwei Beta-Interferon-1a Dosierungsschemata bei RRMS: - Rebif®: 44 µg; s. c., 3x pro Woche - Avonex®: 30 µg; i. m., 1x pro Woche Primärer Endpunkt: Anteil schubfreier Patienten über 24 Wochen. MRT-Endpunkt: Anzahl der aktiven Läsionen pro Patienten-Scan nach 24 Wochen.
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisiert, kontrolliert, multizentrisch Randomisierungsverhältnis: 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine Angaben*.
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: - Interferon-naïve Patienten (18 – 55 Jahre) mit nachgewiesener RRMS; - Mindestens 2 Schübe in den letzten 2 Jahren; - EDSS-Score von 0 bis 5,5 Ausschlusskriterien: - Vorherige Verwendung von Interferon oder Cladribin; - Totale Lymphbestrahlung; - Glatirameracetat- oder Zytokin-Therapie in den vergangenen 3 Monaten; - Immunglobulin (intravenös) in den vergangenen 6 Monaten; - andere immunmodulatorische Agentien in den vergangenen 12 Monaten;
4b	Studienorganisation und Ort der Studiendurchführung	56 Zentren (15 in Europa, 5 in Kanada, 36 in USA)
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Prüfsubstanzen: Beta-Interferon-1a - Rebif®: 44 µg; s. c., 3x pro Woche; Verabreichung gemäß Empfehlung des Herstellers. - Avonex®: 30 µg; i. m., 1x pro Woche; Als flüssige Formulierung in vorgefüllten Spritzen (dem behandelnden Arzt wurden Richtlinien, basierend auf der WHO-Skala zum Nebenwirkungsschweregrad, für Dosisreduktion

Item ^a	Charakteristikum	Studieninformation
		oder –unterbrechung zur Verfügung gestellt).
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Endpunkt: Anteil der für 24 Wochen schubfreien Patienten, wobei das primäre Ziel ein Wahrscheinlichkeitsverhältnis ist (angepasst je Studienzentrum).</p> <p>Sekundäre und weitere Zielparame-ter: Schub-Rate (Schübe pro Patient während der Studie), Schub-Schwere, Steroid-Gebrauch bei Schüben, Zeit bis zum ersten Schubs.</p> <p>Behinderung wurde als Progression um einen Punkt auf der EDSS-Skala definiert, wobei dies bei einer Visite, 3 oder 6 Monate später bestätigt werden musste (in der Zwischenzeit kein EDSS-Wert, der nicht dem Progressionskriterium entsprechen würde).</p> <p>Primäre MRT-Zielparame-ter: Anzahl der aktiven CU (combined unique)-Läsionen pro Patient pro Scan. Eine aktive T2-Läsion wurde als neue oder sich vergrößernde Läsion oder als eine Läsion, die an einer Stelle wiederaufgetreten ist, an welcher sich frühere Läsionen aufgelöst hatten, definiert. Eine CU-Läsion wurde definiert als eine aktive Läsion in T1 post-Gadolinium- oder T2-Sequenzen, oder in beiden Sequenzen (unter Vermeidung von Doppelzählungen).</p> <p>Sekundäre MRT-Zielparame-ter: Anzahl der T2- und T1-Läsionen pro Patient pro Scan, Anteil der aktiven Scans (T2, T1 und CU) pro Patient, Anteil an Patienten- mit und ohne aktive Scans (T2, T1 und CU) während der ersten 24 Wochen der Studie. Bei Woche 48 wurden nur die T2 Läsionen gezählt und gemessen, da Gadolinium nicht verabreicht wurde.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine Angaben*
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Die Fallzahlen wurden über einen 2-seitigen χ^2-Test unter Annahme eines Fehlers der 1. Art von 5% bestimmt. Ferner wurde von 65 % schubfreien Patienten nach 24 Wochen in der Gruppe mit 44 µg 3x pro Woche sowie von 50% in der Gruppe mit 30 µg 1x pro Woche ausgegangen. (dabei wurden Daten aus früheren Studien berücksichtigt).</p> <p>Eine Fallzahl von 280 auswertbaren Patienten pro Behandlungsarm ergab eine 95 %-ige Power zur Detektion eines 30 %-igen Anstiegs im primären Endpunkt bei 24 Wochen in der Gruppe mit 44 µg 3x pro Woche im Vergleich zur Gruppe mit 30 µg</p>

Item ^a	Charakteristikum	Studieninformation
		1x pro Woche. Unter der Annahme von 10 % abbrechenden/nicht auswertbaren Patienten wurden 312 Patienten pro Gruppe benötigt.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Keine Angaben*
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Computer-generierte Randomisierungsliste
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Es gab eine Stratifizierung pro Zentrum mit einem initialen 6er-Block und darauffolgenden 4-er Blöcken, um die Möglichkeiten der Zentren zu verringern, eine nachfolgende Behandlungszuteilung auf Basis früherer Zuteilungen zu bestimmen.
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	zentralisiertes, telefonisches Randomisierungssystem
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Zuteilung zu den Behandlungsgruppen wurde dem unverblindeten Zentrumspersonal über ein zentralisiertes, telefonisches Randomisierungssystem mitgeteilt.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	<p>a) Patienten waren nicht verblindet. b) Behandeltes Personal war nicht verblindet. c) Evaluierende Ärzte (neurologische und MRT Untersuchungen) waren verblindet.</p> <p>Vor Randomisierung wurden behandelnde und evaluierende Ärzte bestimmt. Patienten wurden angewiesen, ihre Behandlungszuordnung oder Symptome (mit Bezug zu ihrer Behandlung), nicht den verblindeten, evaluierenden Ärzten mitzuteilen sowie ihre Injektionsstellen zu verbergen. Die Kommunikation der evaluierenden Ärzte mit den Patienten war auf die benötigten Angaben beschränkt. Die MRT-Auswertung wurde zentral durch verblindetes Personal, welches keine Kenntnis über die Behandlung oder das Behandlungsergebnis der Patienten hatte, durchgeführt. Ein Pflicht-Multiple-Choice-Fragebogen hinsichtlich der Studienverblindung wurde nach Woche 48 an die evaluierenden Ärzte verteilt.</p>

Item ^a	Charakteristikum	Studieninformation
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Auf eine Doppel-Dummy-Lösung, bei der jeder Patient den gleichen Dosierungsplan verwendet und sich sowohl Placebo als auch aktive Prüfsubstanz injiziert, wurde durch die Prüfarzte bewusst verzichtet.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p><u>Primärer Endpunkt</u>: logistische Regression mit Anpassung an Behandlung und Zentrum;</p> <p><u>Das Wahrscheinlichkeitsverhältnis (OR)</u> wurde definiert als Verhältnis der Wahrscheinlichkeit der Schub-Freiheit unter Beta-Interferon-1a 44 µg 3x pro Woche (Anteil der Schub-freien Patienten/Anteil der Patienten mit Schub) dividiert durch die Wahrscheinlichkeit der Schub-Freiheit unter Beta-Interferon-1a 30 µg 1x pro Woche;</p> <p><u>Dauer bis zur bestätigten Progression</u>: Kaplan-Meier-Schätzer wurden berechnet;</p> <p><u>Risiko-Quotienten (HRs), assoziierte Konfidenzintervalle und p-Werte der Behandlungsvergleiche</u>: zu Cox- proportionales Risiko-Modell mit Faktoren für Behandlung und Zentrum;</p> <p><u>Schub-Rate</u> (basierend auf allen berichteten Schüben): Poisson-Regressionsmodell mit Faktoren für Behandlung und Zentrum;</p> <p><u>Anteil an Patienten- ohne CU-, T2- und T1-aktive Läsionen</u>: logistisches Regressionsmodell (angepasst an Behandlung und Zentrum);</p> <p>Alle anderen <u>MRT-Daten</u>: nichtparametrische Kovarianzanalyse (ANCOVA) (angepasst an Behandlung und Zentrum) mit der Anzahl von aktiven Baseline-Läsionen als alleinige Kovariate.</p> <p><u>Verschlechterungsanteil</u>: Poisson-Regressionsmodell mit Faktoren für Behandlung und Zentrum und dem Logarithmus des Studienzeitpunkts als Offset-Variable.</p> <p><u>Risiko-Quotient und assoziiertes 95%-iges Konfidenzintervall bis zur ersten Verschlechterung</u>: zu Cox proportionales Risiko-Modell;</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Bildung von neutralisierenden Antikörpern in beiden Behandlungsarmen.
-	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	Beta-Interferon-1a (44 µg, s. c., 3 x pro Woche): a) 339 Patienten; b) 339 Patienten; c) 339 Patienten;

Item ^a	Charakteristikum	Studieninformation
	Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Beta-Interferon-1a (30 µg, i. m., einmal pro Woche): a) 338 Patienten; b) 337 Patienten; c) 338 Patienten;
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Beta-Interferon-1a (44 µg, s. c., 3 x pro Woche): Abbrüche insgesamt: n = 25 16 wegen unerwünschter Ereignisse 3 wegen Entscheidung der Patienten 3 wegen mangelnder Wirksamkeit 2 wegen Schwangerschaft 1 Unfalltod Beta-Interferon-1a (30 µg, i. m., einmal pro Woche): Abbrüche insgesamt: n = 21 14 wegen unerwünschter Ereignisse 4 wegen Entscheidung der Patienten 1 wegen mangelnder Wirksamkeit 1 ohne aktuelle Verlaufskontrolle 1 vor dem ersten Studientag
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Screening fand 4 Wochen vor Randomisierung und vor Beginn der Studientherapie statt. Ein Follow-up wurde bei Woche 48 durchgeführt. Keine weiteren Angaben *.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde wie geplant abgeschlossen.
a: nach CONSORT 2010.		

* Keine Angaben hierzu in der Publikation.

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

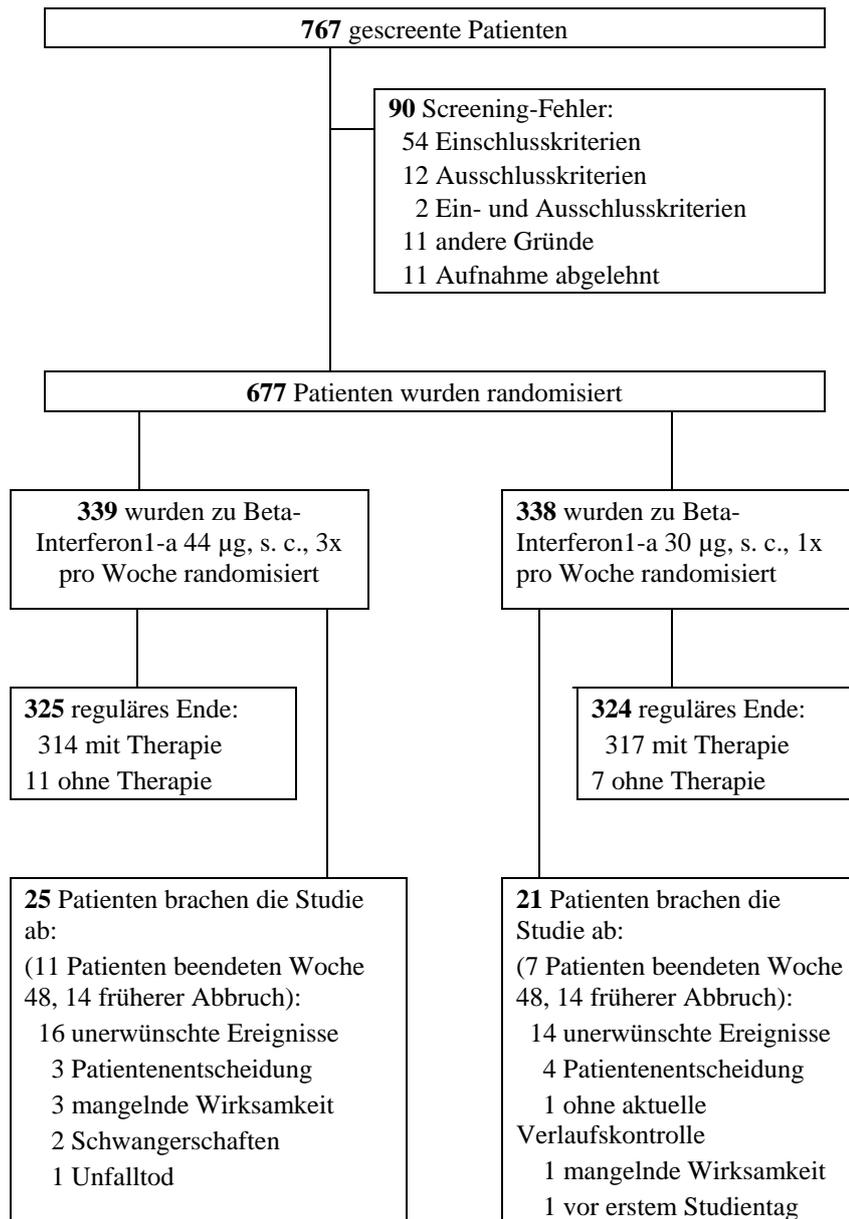


Tabelle 4-99 (Anhang): Studiendesign und -methodik für Studie PRISMS

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	<p>Fragestellung: Wie ist der Grad der Wirkung von Interferon-beta bei MS?</p> <p>Hypothese: Interferon-beta-1a senkt die Schubrate.</p> <p>Primärziel: Bestimmung der Schubrate im Verlauf der Studie</p> <p>Sekundärziele:</p> <ul style="list-style-type: none"> - Zeit bis zum ersten und bis zum zweiten Schub - Anteil schubfreier Patienten - Behinderungsprogression (EDSS-Erhöhung um mindestens 1 Punkt über mindestens 3 Monate) - Erfassung eines Ambulations-Indexes - Erfassung eines Arm-Funktions-Indexes - Bedarf einer Steroidtherapie - Hospitalisierungsrate - Krankheitsverlauf im MRT - Krankheitslast
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	randomisiert, doppelblind, Placebo-kontrolliert, parallel
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Einschlusskriterien:</p> <ul style="list-style-type: none"> - klinisch definierte oder Labor-unterstützte Diagnose von MS seit mindestens einem Jahr - erwachsene RRMS-Patienten - mindestens zwei Schübe innerhalb der letzten zwei Jahre vor Einschluss - Kurtzke EDSS Score 0 – 5,0 <p>Ausschlusskriterien:</p> <ul style="list-style-type: none"> - jegliche vorhergehende systemische Behandlung mit Interferon <p>Innerhalb der letzten zwölf Monate vor Einschluss:</p> <ul style="list-style-type: none"> - Bestrahlung des Lymphsystems - Cyclophosphamid-Behandlung,

Item ^a	Charakteristikum	Studieninformation
		- oder Behandlung mit anderen Immunmodulatoren oder Immunsuppressiva
4b	Studienorganisation und Ort der Studiendurchführung	<p>22 Zentren weltweit in neun Ländern: Australien, Belgien, Deutschland, Finnland, Kanada, Niederlande, Schweden, Schweiz, vereinigtes Königreich</p> <p>Unter Leitung des Sponsors und eines koordinierenden Investigator Liaison Committees wurde ein Independend Monitoring Committee zur Bewertung der Sicherheit und zur Überwachung der Studie eingesetzt.</p>
5	<p>Interventionen</p> <p>Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.</p>	<p>Prüfsubstanz:</p> <p>Interferon-beta-1a 22µg: 22 µg Interferon-beta-1a (6 Millionen IE) in 0,5 ml Lösung zur s.c. Selbstinjektion.</p> <p>Interferon-beta-1a 44µg: 44 µg Interferon-beta-1a (12 Millionen IE) in 0,5 ml Lösung zur s.c. Selbstinjektion.</p> <p>Die Interferon-Dosierungen wurden über 4-8 Wochen graduell erhöht; für 2-4 Wochen mit 20 % der finalen Dosis und für weitere 2-4 Wochen mit 50 % der finalen Dosis und anschließend der Zieldosis. Bei Auftreten von Toxizitäten WHO-Grad II oder III geschah eine Dosisanpassung auf 50 % oder eine zeitweise Aussetzung der Injektion.</p> <p>Kontrollsubstanz:</p> <p>Placebo: 0,5 ml Lösung zur s.c. Selbstinjektion.</p>
6	Zielkriterien	
6a	<p>Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten</p>	<p>Primärzielkriterien: Bestimmung der Schubraten im Verlauf der Studie, Schub definiert als: Auftreten eines neuen Symptoms oder Verschlechterung eines vorherbestehenden für mindestens 24 Stunden, das einer MS-Aktivität zugeordnet werden kann und zwar nach einer 30-tägigen Phase der Stabilität oder Verbesserung Die Schwere eines Schubes wurde mittels der „Scribbs neurological rating scale“ oder der „activities of daily living scale“ bewertet.</p> <p>Sekundärzielkriterien: Die Erfassung der sekundären Zielparameter geschah vierteljährlich mittels Untersuchungen und anamnetischen Erfassung über die keine Information dargelegt ist .</p> <p>MRT-Untersuchungen wurden zentral und verblindet ausgewertet</p>

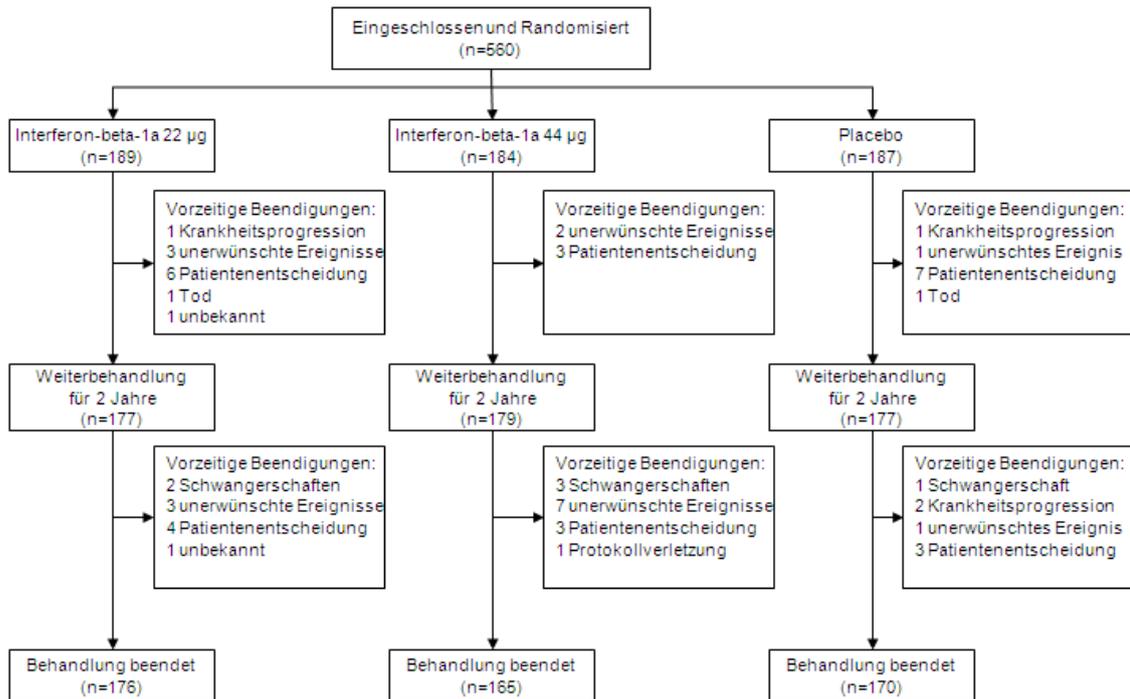
Item ^a	Charakteristikum	Studieninformation
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Keine Angaben*
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Zur Bewertung der Sicherheit wurde ein unabhängiges Monitoring Komitee installiert. Es wurden keine Zwischenanalysen durchgeführt.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Computergenerierte Randomisierungslisten von Serono Biometrics zur Verwendung an den Zentren.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Blockrandomisierung mit Blockgröße 6, Stratifizierung nach Zentren.
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	Computer-generierte Liste zur Zuteilung an den Zentren. Ansonsten keine Information*. Keine Berichte von Entblindungen.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Zuteilung erfolgte anhand einer Computergenerierten Liste an den Zentren. Ansonsten keine weitere Information*
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) Patienten waren verblindet, die Verpackung und Aussehen der Präparate war neutral und untereinander identisch. b) Jegliches beteiligte Personal war verblindet. Ein „Behandlungs-“Arzt war für das Medikationsmanagement verantwortlich, inklusive der Behandlung und Erfassung von Nebenwirkungen. c) zur Beurteilung von Schüben und für neurologische Untersuchungen wurde ein „Untersuchungs-“Arzt eingesetzt. Bei den neurologischen Untersuchungen wurden die Injektionsstellen abgedeckt damit durch Hautirritationen keine Gefahr der Entblindung bestand.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die drei eingesetzten Präparate glichen sich in Aussehen (neutrale und identische Verpackung) und Applikationsart (Selbstinjektion von 0,5 ml s.c.).
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Analyse wurde mit der Intent-to-Treat Population durchgeführt. Alle Daten wurden

Item ^a	Charakteristikum	Studieninformation
		<p>berücksichtigt. Die Studie hatte eine Power von 80 %, einen Unterschied von 0,64 der mittleren Schubhäufigkeit aufzuzeigen (bezogen auf Placebo vs. 44 µg Interferon-beta-1a). Mit einer Effektgröße von 0,4 (MW 0,64; SD 1,62) wurde für einen zweiseitigen Zwei-Stichproben-<i>t</i>-Test ($p=0,05$) ein Minimum von 100 Patienten pro Gruppe benötigt. Ein generalisiertes lineares Model (GLM) mit log Link-Funktion und Varianzen proportional zum MW wurde eingesetzt, um die Schubhäufigkeit zu berechnen. Es wurde Cox proportionales Risiko Modell für „Zeit bis Eintreten“ Endpunkte sowie logistische Regression für binäre Endpunkte eingesetzt, ANCOVA über Rang-Daten für weitere kontinuierliche Endpunkte und Chi-QuadratTests für Patientenzahlen zu speziellen Kategorien unerwünschter Ereignisse.</p> <p>Laut Prüfplan lag der Fokus auf Vergleichen der Placebo und der 44 µg Interferon-beta-1a Gruppen. Es wurden keine Trendanalysen für alle drei Behandlungsgruppen durchgeführt.</p> <p>Für die Analyse der Untergruppe stärker vorgeschädigter Patienten ($EDSS \geq 3,5$) wurde ein Modell eingesetzt, welches nach Behandlung, Zentren, Baseline EDSS Kohorte und Behandlung nach Baseline-Kohorte unterschied.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Es erfolgten Subgruppenanalysen für Patienten mit schwerer Behinderung bei Einschluss ($EDSS \geq 3,5$). Außerdem wurden 267 Patienten hinsichtlich Depression untersucht. Die Analyse zur Behinderungsprogression (EDSS) wurde adjustiert, hier wurde der IDSS (integrated disability status scale) bestimmt, die AUC (Area under the curve) des EDSS.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<p>Interferon-beta-1a 22µg:</p> <p>a) 189 b) 189* c) 189</p> <p>Interferon-beta-1a 44µg:</p> <p>a) 184 b) 184* c) 184</p> <p>Placebo:</p> <p>a) 187 b) 187* c) 187</p>

Item ^a	Charakteristikum	Studieninformation
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Interferon-beta-1a 22µg: nach 1 Jahr (12): 1 Krankheitsprogression, 3 AE, 6 Patientenentscheidung, 1 Tod, 1 unbekannt nach 2 Jahren (10): 2 Schwangerschaft, 3 AE, 4 Patientenentscheidung, 1 unbekannt</p> <p>Interferon-beta-1a 44µg: nach 1 Jahr (5): 2 AE, 3 Patientenentscheidung nach 2 Jahren (14): 3 Schwangerschaft, 7 AE, 3 Patientenentscheidung, 1 Protokollverletzung</p> <p>Placebo: nach 1 Jahr (10): 1 Krankheitsprogression, 1 AE, 7 Patientenentscheidung, 1 Tod nach 2 Jahren (7): 1 Schwangerschaft, 2 Krankheitsprogression, 1 AE, 3 Patientenentscheidung</p>
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Patienten wurden zwischen Mai 1994 und Februar 1995 eingeschlossen und zwei Jahre behandelt. Keine Information zu Nachbeobachtungen*.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde regulär beendet.
a: nach CONSORT 2010.		

* Keine Angaben hierzu in der Publikation.

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.



Anhang 4-G: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten

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.

Studie: TRANSFORMS

Tabelle 4-54 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie TRANSFORMS

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** 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 werden die Bögen studienbezogen ausgefüllt. 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.

Tabelle: Liste der für die Bewertung herangezogenen Quellen

4.8 Genaue Benennung der Quelle	4.9 Kürzel
Studienbericht inklusive Studienprotokoll (23)	FTY720D2302

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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 kodierte 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 Anhaltspunkte 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 Behandlern**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:

Alle Patienten waren verblindet und bekamen entweder einmal täglich Fingolimod 0.5 mg, einmal täglich Fingolimod 1.25 mg oder einmal wöchentlich IFN- β -1a i. m. 30 μ g in einem double-dummy Design.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Um die Verblindung zu garantieren wurden die Patienten angewiesen ihre Injektionsstelle vor jeder geplanten Visite und schubbedingter neurologischen Untersuchung abzudecken (z. B. mit einem Pflaster oder dementsprechender Kleidung um die Injektionsstelle komplett zu verdecken) und unerwünschte Ereignisse (z. B. Einstichstellenreaktion oder Grippe ähnliche Symptome) nicht mit dem unabhängigen untersuchenden Arzt zu besprechen.

Der unabhängige untersuchende Arzt war verantwortlich für

- die Bestimmung der EDSS und MSFC-Werte basierend auf einer detaillierten neurologischen Untersuchung wie im Protokoll vorgesehen. Der MSFC-Wert konnte auch von darauf trainiertem Personal am Zentrum, das nicht in die Behandlung des Patienten involviert war, erhoben werden.
- die Bestimmung des EDSS-Wertes zur Schubbestätigung an jeder unplanmäßigen Visite falls vom behandelnden Arzt ein MS-Schub berichtet wurde.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte 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. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

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

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: Jährliche Schubrate über 12 Monate

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

Ein unabhängiger beurteilender Arzt war verantwortlich für:

Bestimmung des EDSS-Wertes zur Schubbestätigung an jeder unplanmäßigen Visite falls vom behandelnden Arzt ein MS-Schub berichtet wurde.

Ein unabhängiger beurteilender Arzt war in keiner Hinsicht in die Behandlung und Management des Patienten involviert.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine

Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Schübe zu Monat 12

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

Ein unabhängiger beurteilender Arzt war verantwortlich für:

Bestimmung des EDSS-Wertes zur Schubbestätigung an jeder unplanmäßigen Visite falls vom behandelnden Arzt ein MS-Schub berichtet wurde.

Ein unabhängiger beurteilender Arzt war in keiner Hinsicht in die Behandlung und Management des Patienten involviert.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anzahl milder, moderater oder schwerer Schübe**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

Ein unabhängiger beurteilender Arzt war verantwortlich für:

Bestimmung des EDSS-Wertes zur Schubbestätigung an jeder unplanmäßigen Visite falls vom behandelnden Arzt ein MS-Schub berichtet wurde.

Ein unabhängiger beurteilender Arzt war in keiner Hinsicht in die Behandlung und Management des Patienten involviert.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schübe, die einen Krankenhausaufenthalt erfordern

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

Ein unabhängiger beurteilender Arzt war verantwortlich für:

Bestimmung des EDSS-Wertes zur Schubbestätigung an jeder unplanmäßigen Visite

falls vom behandelnden Arzt ein MS-Schub berichtet wurde.
Ein unabhängiger beurteilender Arzt war in keiner Hinsicht in die Behandlung und Management des Patienten involviert.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schübe, die eine Steroidbehandlung erfordern

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

Ein unabhängiger beurteilender Arzt war verantwortlich für:

Bestimmung des EDSS-Wertes zur Schubbestätigung an jeder unplanmäßigen Visite falls vom behandelnden Arzt ein MS-Schub berichtet wurde.

Ein unabhängiger beurteilender Arzt war in keiner Hinsicht in die Behandlung und Management des Patienten involviert.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler,

Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Behinderungsprogression, gemessen an der EDSS, über 12 Monate**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 EDSS-Bewertungen wurden von einem unabhängigen beurteilenden Arzt, der nicht in die Behandlung der Patienten involviert war, ausgeführt um eine potenzielle Verzerrungen aufgrund der möglichen Entblindung von unerwünschten Ereignissen zu verhindern.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten

ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5%) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Mittlere Änderung der EDSS über 12 Monate

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 EDSS-Bewertungen wurden von einem unabhängigen beurteilenden Arzt, der nicht in die Behandlung der Patienten involviert war, ausgeführt um eine potenzielle Verzerrungen aufgrund der möglichen Entblindung von unerwünschten Ereignissen zu verhindern.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Mittlere Änderung des Multiple Sclerosis Functional Composite (MSFC) über 12 Monate

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

MSFC-Bewertung wurde entweder von einem unabhängigen beurteilenden Arzt oder einer qualifizierten Person durchgeführt, die nicht in die Behandlung der Patienten involviert war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Mittlere Änderung des EQ-5D über 12 Monate**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

Die EQ-5D hat 2 Komponenten:

Health State Classification – hier selektiert der verblindete Patient selbst die Antwort die am besten auf seinen momentanen Gesundheitszustand zutrifft.

Visual Analogue Scale ‘Thermometer – der verblindete Patient markiert auf der

Thermometer Skala von 0 (denkbar schlimmste Gesundheitszustand) bis 100 (denkbar beste Gesundheitszustand) seinen momentanen Gesundheitszustand.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse

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

Unerwünschte Ereignisse wurden von dem verblindeten primär behandelnden Arzt oder der verblindeten Studienkrankenschwester erfasst.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler,

Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: schwerwiegende unerwünschte Ereignisse**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

Unerwünschte Ereignisse wurden von dem verblindeten primär behandelnden Arzt oder der verblindeten Studienkrankenschwester erfasst.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“-)Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen

Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5%) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Behandlungsabbrüche aufgrund unerwünschter Ereignisse**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

Unerwünschte Ereignisse wurden von dem verblindeten primär behandelnden Arzt oder der verblindeten Studienkrankenschwester erfasst.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Relevante unerwünschte Ereignisse

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

Unerwünschte Ereignisse wurden von dem verblindeten primär behandelnden Arzt oder der verblindeten Studienkrankenschwester erfasst.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Studie: FREEDOMS

Tabelle 4-100 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie FREEDOMS

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** 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 werden die Bögen studienbezogen ausgefüllt. 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.

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Studienbericht inklusive Studienprotokoll (22)	FTY720D2301

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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 kodierte 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 Anhaltspunkte 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 Behandlern**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:

Alle Patienten waren verblindet und bekamen entweder einmal täglich Fingolimod 0.5 mg, einmal täglich Fingolimod 1.25 mg oder entsprechendes Placebo.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den

primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Alle in der Studie eingesetzten Tabletten (Fingolimod in zwei Dosisstufen und Placebo) waren äußerlich identisch. Alle in die Studie involvierten Personen (Patienten, Personal am Zentrum, Personal beim Sponsor außer DSM und unabhängige Mitglieder des DSMB) waren zwischen Randomisierung bis Ende der Studie verblindet.

Der unabhängige untersuchende Arzt war verantwortlich für

- die Bestimmung der EDSS und MSFC-Werte basierend auf einer detaillierten neurologischen Untersuchung wie im Protokoll vorgesehen. Der MSFC-Wert konnte auch von darauf trainiertem Personal am Zentrum, das nicht in die Behandlung des Patienten involviert war, erhoben werden.
- die Bestimmung des EDSS-Wertes zur Schubbestätigung an jeder unplanmäßigen Visite falls vom behandelnden Arzt ein MS-Schub berichtet wurde.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte 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. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

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

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: Jährliche Schubrate über 24 Monate

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

Ein unabhängiger beurteilender Arzt war verantwortlich für:

Bestimmung des EDSS-Wertes zur Schubbestätigung an jeder unplanmäßigen Visite falls vom behandelnden Arzt ein MS-Schub berichtet wurde.

Ein unabhängiger beurteilender Arzt war in keiner Hinsicht in die Behandlung und Management des Patienten involviert.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrucher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass

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

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5%) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Schübe zu Monat 24**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

Ein unabhängiger beurteilender Arzt war verantwortlich für:

Bestimmung des EDSS-Wertes zur Schubbestätigung an jeder unplanmäßigen Visite falls vom behandelnden Arzt ein MS-Schub berichtet wurde.

Ein unabhängiger beurteilender Arzt war in keiner Hinsicht in die Behandlung und Management des Patienten involviert.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anzahl milder, moderater oder schwerer Schübe

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

Ein unabhängiger beurteilender Arzt war verantwortlich für:

Bestimmung des EDSS-Wertes zur Schubbestätigung an jeder unplanmäßigen Visite falls vom behandelnden Arzt ein MS-Schub berichtet wurde.

Ein unabhängiger beurteilender Arzt war in keiner Hinsicht in die Behandlung und Management des Patienten involviert.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schübe, die einen Krankenhausaufenthalt erfordern

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
Ein unabhängiger beurteilender Arzt war verantwortlich für:

Bestimmung des EDSS-Wertes zur Schubbestätigung an jeder unplanmäßigen Visite falls vom behandelnden Arzt ein MS-Schub berichtet wurde.

Ein unabhängiger beurteilender Arzt war in keiner Hinsicht in die Behandlung und Management des Patienten involviert.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schübe, die eine Steroidbehandlung erfordern

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

Ein unabhängiger beurteilender Arzt war verantwortlich für:

Bestimmung des EDSS-Wertes zur Schubbestätigung an jeder unplanmäßigen Visite falls vom behandelnden Arzt ein MS-Schub berichtet wurde.

Ein unabhängiger beurteilender Arzt war in keiner Hinsicht in die Behandlung und Management des Patienten involviert.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler,

Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Behinderungsprogression, gemessen an der EDSS, über 12 Monate**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 EDSS-Bewertungen wurden von einem unabhängigen beurteilenden Arzt, der nicht in die Behandlung der Patienten involviert war, ausgeführt um eine potenzielle Verzerrungen aufgrund der möglichen Entblindung von unerwünschten Ereignissen zu verhindern.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Mittlere Änderung der EDSS über 12 Monate

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 EDSS-Bewertungen wurden von einem unabhängigen beurteilenden Arzt, der nicht in die Behandlung der Patienten involviert war, ausgeführt um eine potenzielle Verzerrungen aufgrund der möglichen Entblindung von unerwünschten Ereignissen zu verhindern.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Mittlere Änderung des Multiple Sclerosis Functional Composite (MSFC) über 12 Monate

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

MSFC-Bewertung wurde entweder von einem unabhängigen beurteilenden Arzt oder einer qualifizierten Person durchgeführt, die nicht in die Behandlung der Patienten involviert war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5%) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Mittlere Änderung des EQ-5D über 12 Monate**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

Die EQ-5D hat 2 Komponenten:

Health State Classification – hier selektiert der verblindete Patient selbst die Antwort die am besten auf seinen momentanen Gesundheitszustand zutrifft.

Visual Analogue Scale ‘Thermometer – der verblindete Patient markiert auf der

Thermometer Skala von 0 (denkbar schlimmste Gesundheitszustand) bis 100 (denkbar beste Gesundheitszustand) seinen momentanen Gesundheitszustand.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse

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

Unerwünschte Ereignisse wurden von dem verblindeten primär behandelnden Arzt oder der verblindeten Studienkrankenschwester erfasst.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler,

Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: schwerwiegende unerwünschte Ereignisse**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

Unerwünschte Ereignisse wurden von dem verblindeten primär behandelnden Arzt oder der verblindeten Studienkrankenschwester erfasst.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen

Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5%) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Behandlungsabbrüche aufgrund unerwünschter Ereignisse**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

Unerwünschte Ereignisse wurden von dem verblindeten primär behandelnden Arzt oder der verblindeten Studienkrankenschwester erfasst.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Relevante unerwünschte Ereignisse

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

Unerwünschte Ereignisse wurden von dem verblindeten primär behandelnden Arzt oder der verblindeten Studienkrankenschwester erfasst.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Studie Bornstein et al. (1987)

Tabelle 4-56 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Bornstein et al., 1987

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** 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 werden die Bögen studienbezogen ausgefüllt. 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.

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Manuskript: Bornstein MB, Miller A. Slagle S et al. A pilot trial of Cop 1 in exacerbating-remitting multiple sclerosis. N Engl J Med. 1987; 317(7):408-414 (74)	Bornstein et al.

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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:

In dieser matched pair Studie wurde die Zuteilung des matched Patienten bestimmt durch die Randomisierung des ersten Patienten.

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:

In dieser matched pair Studie wurde die Zuteilung des matched Patienten bestimmt durch die Randomisierung des ersten Patienten.

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 Anhaltspunkte 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 Behandlern**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:

Alle Patienten waren verblindet und mußten sich spritzen entweder mit 20 mg Cop 1 aufgelöst in 1ml Salzlösung oder nur Salzlösung einmal täglich für 2 Jahre.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Ein verblindeter Neurologe machte an jeder Visite die neurologischen Untersuchungen und Status Einschätzungen. Die Selbsteinschätzung des Patienten bezüglich Nebenwirkungen und Veränderungen im Neurologischen Status wurde dem klinischen Assistenten berichtet, der nicht verblindet war. Der Neurologe und Patient vermieden es über Nebenwirkungen zu reden.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte 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. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

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

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: Jährliche Schubrate über 24 Monate**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

Der verblindete Neurologe verifizierte den Schub, basierend auf den Studienkriterien.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrucher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Schübe zu Monat 24**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
Der verblindete Neurologe verifizierte den Schub, basierend auf den Studienkriterien.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

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

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Behinderungsprogression, gemessen an der EDSS, über 24 Monate**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
EDSS-Bewertungen wurden von einem verblindeten Neurologen durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Studie: Johnson et al., (1995)

Tabelle 4-57 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Johnson et al., 1995

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** 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 werden die Bögen studienbezogen ausgefüllt. 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.

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Manuskript: Johnson KP, Brooks BR, Cohen JA et al. Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: results of a phase III multicenter, double-blind placebo-controlled trial. The Copolymer 1 Multiple Sclerosis Study Group. Neurology. 1995;45(7):1268-1276 (67)	Johnson et al.

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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:

Ein zentralisiertes Randomisierungsschema wurde angewendet.

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 kodierte 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:

Ein zentralisiertes Randomisierungsschema wurde angewendet.

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 Anhaltspunkte 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 Behandlern**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:

Alle Patienten waren verblindet und mußten sich spritzen entweder mit 20 mg Cop 1 aufgelöst in 1ml Salzlösung oder nur Salzlösung einmal täglich für 2 Jahre.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:
Ein verblindeter Neurologe machte an jeder Visite die neurologischen Untersuchungen und Status Einschätzungen. Die Selbsteinschätzung des Patienten bezüglich Nebenwirkungen und Veränderungen im Neurologischen Status wurde dem klinischen Assistenten berichtet, der nicht verblindet war. Der Neurologe und Patient vermieden es über Nebenwirkungen zu reden.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte 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. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

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

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: Jährliche Schubrate über 24 Monate**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

Der verblindete Neurologe verifizierte den Schub, basierend auf den Studienkriterien.

.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Schübe zu Monat 24**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

Der verblindete Neurologe verifizierte den Schub, basierend auf den Studienkriterien.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrucher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Behinderungsprogression, gemessen an der EDSS, über 24 Monate

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 EDSS-Bewertungen wurden von einem verblindeten Neurologen durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Studie: The IFNB Multiple Sclerosis Study Group

Tabelle 4-58 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie der IFNB Multiple Sclerosis Study Group, 1993

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** 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 werden die Bögen studienbezogen ausgefüllt. 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.

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Manuskript: The IFNB Multiple Sclerosis Study Group. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. I. Clinical results of a multicenter, randomized, double-blind, placebo-controlled trial. Neurology 1993;43:655-661. (39)	The IFNB Multiple Sclerosis Study Group

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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:

Ein zentralisiertes Randomisierungsschema wurde angewendet.

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 Anhaltspunkte 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 Behandlern**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:

Alle Patienten waren verblindet und mußten sich spritzen entweder mit IFNB 1.6 MIU, IFNB 8 MIU oder Placebo (albumin und Dextrose) jeden 2. Tag für 2 Jahre.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

2 Ärzte in jedem Zentrum wurden für die Studie bestimmt: Ein Neurologe der die Studienvisiten durchführte aber nicht unerwünschte Ereignisse erfasste, und ein behandelnder Neurologe der die unerwünschten Ereignisse erfasste. Alle Personen am Zentrum waren verblindet.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte 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. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

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

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: Jährliche Schubrate über 24 Monate**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

Der verblindete behandelnde Neurologe verifizierte den Schub, basierend auf den Studienkriterien.

.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Schübe zu Monat 24**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

Der verblindete behandelnde Neurologe verifizierte den Schub, basierend auf den Studienkriterien.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Behinderungsprogression, gemessen an der EDSS, über 36 Monate

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 EDSS-Bewertungen wurden von einem verblindeten Studien Neurologen durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5%) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Behandlungsabbrüche aufgrund unerwünschter Ereignisse über 24 Monate

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

Unerwünschte Ereignisse wurden von einem verblindeten behandelnden Neurologen erfasst.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass

diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Studie: MSCRG

Tabelle 4-59 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie der Multiple Sclerosis Collaborative Research Group, 1996

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** 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 werden die Bögen studienbezogen ausgefüllt. 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.

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Manuskript: Jacobs et al. and The Multiple Sclerosis Collaborative Research Group. Intramuscular Interferon beta-1a for disease progression in relapsing multiple sclerosis. Ann neurol 1996;39:285-294. (41)	MSCRG

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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 kodierte 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:

Die Efron's biased coin method wurde für die Randomisierung verwendet.
Angaben zur Verwendung von identisch aussehenden, nummerierten oder kodierte Arzneimittelbehältern fehlen.

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 Anhaltspunkte 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 Behandlern**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:

Alle Patienten waren verblindet und besprachen keine medizinischen Angelegenheiten mit den unabhängigen Ärzten, die den EDSS Score ermittelten. Die Injektionen (30 ug) wurden wöchentlich von der Study nurse oder angewiesenem lokalem Personal bis zu 104 Wochen verabreicht. 24 Stunden vor und nach den Injektionen wurde außerdem 650 mg Acetaminophen verabreicht.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der behandelnde, verblindete Arzt evaluierte die Laborresultate, untersuchte den Patienten, und traf alle medizinischen Entscheidungen. Ein unabhängiger, verblindeter Studienarzt ermittelte den EDSS.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte 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. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

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

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: Jährliche Schubrate über 24 Monate

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
Falls der verblindete Studienarzt den Patienten zur Zeit des Schubes nicht untersuchen konnte wurde die Untersuchung auf Band aufgenommen gemäß eines standardisierten Skripts und danach vom Studienarzt bewertet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Schübe zu Monat 24**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
 Falls der verblindete Studienarzt den Patienten zur Zeit des Schubes nicht untersuchen konnte wurde die Untersuchung auf Band aufgenommen gemäß eines standardisierten Skripts und danach vom Studienarzt bewertet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Studie: PRISMS

Tabelle 4-60 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie der PRISMS Study Group, 1998

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** 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 werden die Bögen studienbezogen ausgefüllt. 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.

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Manuskript: The PRISMS Study Group. Randomised double-blind placebo-controlled study of interferon beta-1a in relapsing/remitting multiple sclerosis. Lancet 1998; 352:1498-504. (40)	PRISMS

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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:

Es fand eine computergenerierte Randomisierung statt.

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 kodierte 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:

Es fand eine computergenerierte Randomisierung statt. Angaben zur Verwendung von identisch aussehenden, nummerierten oder kodierte Arzneimittelbehältern sind vorhanden.

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 Anhaltspunkte 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 Behandlern**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:

Das Personal am Zentrum und die Patienten waren verblindet.
Interferon beta-1a (Rebif, Ares-Serono) 22ug, oder 44 µg oder Placebo wurde dreimal wöchentlich als Injektion verabreicht.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Alle an der Studie beteiligten Personen waren verblindet.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte 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. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

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

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: Jährliche Schubrate über 24 Monate**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

Ein behandelnder Neurologe war verantwortlich für die generelle Behandlung des Patienten, was auch die Behandlung unerwünschter Ereignisse einschloss. Ein ‚beurteilender‘ Neurologe war verantwortlich für die neurologische Beurteilung und Nachbetreuung der Schübe. Alle Einstichstellen wurden verdeckt bei der neurologischen Untersuchung um abzusichern, daß die Verblindung nicht kompromittiert wurde wegen lokalen Reaktionen.

.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung

angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Schübe zu Monat 24

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

Ein behandelnder Neurologe war verantwortlich für die generelle Behandlung des Patienten, was auch die Behandlung unerwünschter Ereignisse einschloss. Ein ‚beurteilender‘ Neurologe war verantwortlich für die neurologische Beurteilung und Nachbetreuung der Schübe. Alle Einstichstellen wurden verdeckt bei der neurologischen Untersuchung um abzusichern, daß die Verblindung nicht kompromittiert wurde wegen lokalen Reaktionen.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- un plausible 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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Behandlungsabbrüche aufgrund unerwünschter über 24 Monate

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
Unerwünschte Ereignisse wurden von einem verblindeten behandelnden Neurologen erfasst.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Studie: EVIDENCE

Tabelle 4-61 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie der EVIDENCE Group und der Universität der British Columbia MS/MRI Research Group, 2002

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** 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 werden die Bögen studienbezogen ausgefüllt. 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.

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Manuskript: Panitch et al. for the The EVIDENCE Group and the University of British Columbia. Randomized, comparative study of interferon beta-1a treatment regimens in MS. Neurology 2002;59:1496-1506. (75)	EVIDENCE

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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:

Eine computergenerierte Randomliste wurde erstellt. Die Gruppenzuteilung erfolgte über ein zentrales Telefonrandomisierungssystem.

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:

Ein zentralisiertes Randomisierungsschema wurde angewendet. Die Patienten konnten nicht verblindet werden.

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 Anhaltspunkte 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 Behandlern**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:

Die Patienten waren unverblindet und mußten Ihre 24 wöchige Behandlung mit Rebif[®] 44µg oder mit Avonex[®] 30µg vor verblindetem Personal geheim halten.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Vor der Randomisierung wurden pro Zentrum ein Arzt bestimmt, der verblindet die Studienvisiten durchführt und ein Arzt, der unverblindet für die Behandlung zuständig ist. Bedingt durch die unterschiedliche Medikationsdarreichungsform der beiden Medikamente konnten die Patienten und die behandelnden Ärzte nicht verblindet werden.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte 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. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

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

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: Jährliche Schubrate über 12 Monate

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

Der verblindete Studienneurologe verifizierte den Schub, basierend auf den Studienkriterien.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“-)Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Schübe zu Monat 12

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
Der verblindete Studienneurologe verifizierte den Schub, basierend auf den Studienkriterien.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Behandlungsabbrüche aufgrund unerwünschter über 12 Monate**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
Unerwünschte Ereignisse wurden vom verblindeten Studienneurologen erfasst.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Studie: INCOMIN

Tabelle 4-62 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie der INCOMING Trial Study, 2002

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** 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 werden die Bögen studienbezogen ausgefüllt. 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.

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Manuskript: Durelli L, Verdun E, Barbero P et al. Every-other-day interferon beta-1b versus once-weekly interferon beta-1a for multiple sclerosis: results of a 2-year prospective randomised multicentre study (INCOMIN). Lancet 2002;359:1453-60. (43)	INCOMIN

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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:

Erstellung einer computergenerierten Randomisierungsliste.

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:

Ein Randomisierungsschema wurde angewendet. Die Patienten konnten nicht verblindet werden.

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 Anhaltspunkte 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 Behandlern**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:

Die Patienten waren unverblindet und mussten entweder 250µg interferon beta1-b jeden zweiten Tag subkutan spritzen oder 30µg interferon beta1-b intramuskulär.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Alle klinischen Endpunkte wurden unverblindet erhoben. Die Auswertung der MRI Analyse fand extern statt.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte 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. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

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

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Die klinischen Endpunkte wurden in einem offenen Studiendesign erhoben.

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: Jährliche Schubrate über 24 Monate

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
Alle klinischen Endpunkte wurden unverblindet erhoben.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Die klinischen Endpunkte wurden unverblindet erhoben.

Endpunkt: Anteil Patienten ohne Schübe zu Monat 24

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
Alle klinischen Endpunkte wurden unverblindet erhoben.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:
Alle klinischen Endpunkte wurden unverblindet erhoben.

Endpunkt: Behandlungsabbrüche aufgrund unerwünschter Ereignisse über 12 Monate**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
Klinische Endpunkte wurden unverblindet erfasst.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:
Klinische Endpunkte wurden unverblindet erfasst.

Studie: BEYOND

Tabelle 4-63 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie der BEYOND Study Group, 2009

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** 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 werden die Bögen studienbezogen ausgefüllt. 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.

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Manuskript: O'Connor P, Filippi M, Arnason B et al. 250 µg or 500 µg interferon beta-1b versus 20 mg glatiramer acetate in relapsing-remitting multiple sclerosis: a prospective, randomised, multicentre study. <i>Neurology</i> 2009;8:889-97. (80)	BEYOND

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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:

SAS basierte block-Randomisierung mit Stratifizierung.

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 kodierte 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:

Ein Randomisierungsschema wurde angewendet. Die unterschiedlichen Medikationen waren identisch in Beschriftung, Aussehen und Verpackung.

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 Anhaltspunkte 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 Behandlern**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:

Die Ärzte und die Patienten waren verblindet. Die Behandlung erfolgte nach randomisierter 2:2:1 Zuteilung der Medikation mit entweder 250µg oder 500µg interferon beta1-b jeden zweiten Tag subkutan oder 20mg glatiramer acetate subkutan täglich.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Ärzte und die Patienten waren verblindet.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte 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. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

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

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: Jährliche Schubrate über 24 Monate

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

Die Erhebung der neurologischen Endpunkte, wie Schübe erfolgte durch einen verblindeten Studienneurologen.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“-)Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Schübe zu Monat 24

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
Die Erhebung der neurologischen Endpunkte, wie Schübe erfolgte durch einen verblindeten Studienneurologen.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Behinderungsprogression, gemessen an der EDSS, über 24 Monate**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
Die Erhebung der EDSS erfolgte durch einen verblindeten Studienneurologen.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Behandlungsabbrüche aufgrund unerwünschter Ereignisse über 24 Monate

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
Unerwünschte Ereignisse wurden vom unverblindeten behandelnden Neurologen erfasst. Patienten erwähnten keine unerwünschten Ereignisse gegenüber dem verblindeten Studienneurologen.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Studie:REGARD

Tabelle 4-64 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie der REGARD Trial Study, 2008

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** 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 werden die Bögen studienbezogen ausgefüllt. 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.

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Manuskript: Mikol DD, Barkhof F, Chang P et al. Comparison of subcutaneous interferon beta-1a with glatiramer acetate in patients with relapsing multiple sclerosis (the REbif vs Glatiramer Acetate in Relapsing MS Disease [REGARD] study): a multicentre, randomised, parallel, open-label trial. <i>Neurology</i> 2008;7:903-14. (82)	REGARD

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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:

Erstellung einer computergenerierten Randomisierungsliste, stratifiziert bei Zentrum.

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 kodierte 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:

Die Information ist aus der Publikation nicht erkennbar.

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 Anhaltspunkte 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 Behandlern**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:

Die Patienten waren unverblindet und bekamen entweder 44µg interferon beta1-b dreimal wöchentlich subkutan oder täglich 20mg glatiramer acetate subkutan für 96 Wochen.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und der behandelnde Arzt waren nicht verblindet. Der Arzt, der die Endpunkte (z.B. EDSS) erhob, war verblindet.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte 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. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

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

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: Jährliche Schubrate über 24 Monate

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

Der verblindete Neurologe verifizierte den Schub, basierend auf den Studienkriterien, und kommunizierte mit dem Patienten nur wenn erforderlich um die EDSS oder einen Schub zu erfassen .

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-

Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Schübe zu Monat 24

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
Die Erhebung der neurologischen Endpunkte, wie Schübe anhand EDSS und
patienten-berichteter Ereignisse erfolgte durch einen verblindeten
Studienneurologen.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

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

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

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

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

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Behandlungsabbrüche aufgrund unerwünschter Ereignisse über 24 Monate

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

Es ist unklar von der Publikation wer die unerwünschten Ereignisse erfasst hat.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:
Klinische Endpunkte wurden unverblindet erfasst.

Studie: BECOME

Tabelle 4-65 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie der BECOME Trial Study, 2009

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** 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 werden die Bögen studienbezogen ausgefüllt. 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.

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Manuskript: Cadavid D, Wolansky LJ, Skurnick J et al. Efficacy of treatment of MS with IFN beta-1b or glatiramer acetate by monthly brain MRI in the BECOME study. Neurology 2009;72(23):1976-1983. (81)	BECOME

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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:

Die Gruppenzuteilung erfolgte randomisiert. Eine genauere Beschreibung liegt nicht vor.

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 kodierte 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:

Die Gruppenzuteilung erfolgte randomisiert. Eine genauere Beschreibung liegt nicht vor.

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 Anhaltspunkte 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 Behandlern**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:

Eine Verblindung der Patienten war nicht möglich, bedingt durch die charakteristischen Erscheinungen nach der Injektionsgabe der einzelnen Medikamente.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und der behandelnde Arzt waren nicht verblindet. Der Arzt, der die Untersuchungen durchführte, war verblindet.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte 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. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

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

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: Jährliche Schubrate über 24 Monate**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

Der verblindete, untersuchende Neurologe verifizierte den Schub, basierend auf einer Verschlechterung anhand der SNRS oder der EDSS.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*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.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anteil Patienten ohne Schübe zu Monat 24**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

Der verblindete, untersuchende Neurologe verifizierte den Schub, basierend auf einer Verschlechterung anhand der SNRS oder der EDSS.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten 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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Studie: Etemadifar et al., (2006)

Tabelle 4-66 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für die Studie des University of Medical Sciences and Health Services, Isfahan, Iran, 2006

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser 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:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** 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 werden die Bögen studienbezogen ausgefüllt. 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.

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Manuskript: Etemadifar M, Janghorbani M, Shaygannejad V. Comparison of Betaferon, Avonex, and Rebif in treatment of relapsing–remitting multiple sclerosis. Acta Neurol Scand 2006;113:283-287. (83)	

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

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

A: Kriterien des Verzerrungspotenzials der Ergebnisse 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 Anhaltspunkte 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:

Randomisierte und zufällige Zuteilung zur Behandlungsgruppe. Jedoch keine nähere Beschreibung.

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:

Die Gruppenzuteilung erfolgte zufällig, ist aber nicht genauer beschrieben.

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 Anhaltspunkte 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 Behandlern**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:

Die Patienten waren unverblindet und bekamen entweder 250µg Betaferon jeden zweiten Tag subkutan für 24 Monate oder einmal wöchentlich intramuskulär 30µg Avonex für 24 Monate oder subkutan 44µg Rebif dreimal wöchentlich für 24 Monate.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (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 Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Arzt, der die klinische Beurteilung machte, war verblindet.

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 Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials 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 das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte 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. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

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

ja

nein

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

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

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

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

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse 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: Anteil Patienten ohne Schübe zu Monat 24**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

Klinische Beurteilung erfolgte durch einen verblindeten Arzt.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten

ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5%) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

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

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

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

Eine ITT Population wurde nicht definiert.

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 Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

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

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

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.

falls hoch, obligate Begründung für die Einstufung:

Anhang 4-H: Ergebnistabellen Subgruppenanalysen

Die Subgruppe zur Krankheitsschwere war in der Kategorie EDSS >3,5 für die Subpopulation der Patienten mit schnell fortschreitender schwerer RRMS nicht für jeden Endpunkt abbildbar. Entsprechende Tabellen fehlen daher an diesen Stellen.

Novartis

FTY720 HTA

HTA Table 11.1.1
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=129)	108	2.40 (1.276)	0.123	2.29 (1.512)	0.145	-0.12 (0.864)	0.083	0.173		
Placebo (N=120)	92	2.27 (1.217)	0.127	2.22 (1.534)	0.160	-0.05 (0.956)	0.100			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.1.2
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=296)	266	2.21 (1.258)	0.077	2.26 (1.416)	0.087	0.05 (0.880)	0.054	0.008		
Placebo (N=298)	240	2.47 (1.279)	0.083	2.67 (1.480)	0.096	0.20 (0.920)	0.059			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 11.1.3
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=363)	323	1.92 (0.974)	0.054	1.97 (1.236)	0.069	0.05 (0.852)	0.047	0.033		
Placebo (N=346)	280	2.03 (0.937)	0.056	2.17 (1.250)	0.075	0.14 (0.937)	0.056			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 11.1.4
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=62)	51	4.42 (0.569)	0.080	4.15 (1.230)	0.172	-0.27 (0.992)	0.139	0.012		
Placebo (N=72)	52	4.50 (0.542)	0.075	4.56 (1.136)	0.158	0.06 (0.938)	0.130			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 11.1.5
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=148)	134	2.11 (1.455)	0.126	2.01 (1.391)	0.120	-0.09 (0.734)	0.063	0.278		
Interferon beta-1a i.m. (N=139)	127	2.31 (1.213)	0.108	2.31 (1.389)	0.123	-0.00 (0.798)	0.071			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL

Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.1.6
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=281)	260	2.28 (1.256)	0.078	2.20 (1.393)	0.086	-0.08 (0.824)	0.051	0.137		
Interferon beta-1a i.m. (N=292)	250	2.10 (1.269)	0.080	2.12 (1.363)	0.086	0.02 (0.776)	0.049			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL

Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 11.1.7
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=362)	332	1.80 (0.953)	0.052	1.76 (1.097)	0.060	-0.04 (0.751)	0.041			0.215
Interferon beta-1a i.m. (N=371)	326	1.84 (0.995)	0.055	1.86 (1.150)	0.064	0.02 (0.799)	0.044			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for gender.

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL

Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 11.1.8
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=67)	62	4.47 (0.572)	0.073	4.15 (1.057)	0.134	-0.32 (0.963)	0.122	0.055		
Interferon beta-1a i.m. (N=60)	51	4.26 (0.451)	0.063	4.24 (0.783)	0.110	-0.03 (0.674)	0.094			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for gender.

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.2.1
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=19)	14	2.57 (1.222)	0.327	2.39 (1.509)	0.403	-0.18 (0.823)	0.220	0.456		
Placebo (N=14)	9	2.72 (1.121)	0.374	2.89 (1.654)	0.551	0.17 (0.829)	0.276			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 11.2.2
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=38)	34	2.18 (1.022)	0.175	2.31 (1.462)	0.251	0.13 (1.039)	0.178	0.760		
Placebo (N=38)	27	3.07 (1.736)	0.334	2.81 (1.694)	0.326	-0.26 (1.013)	0.195			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 11.2.3
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=49)	43	2.07 (0.904)	0.138	2.12 (1.313)	0.200	0.05 (0.987)	0.151	0.816		
Placebo (N=31)	21	1.83 (0.926)	0.202	1.74 (0.889)	0.194	-0.10 (0.944)	0.206			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.2.4
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=8)	5	4.20 (0.447)	0.200	4.20 (1.440)	0.644	0.00 (1.061)	0.474	0.994		
Placebo (N=21)	15	4.60 (0.604)	0.156	4.37 (1.202)	0.310	-0.23 (1.050)	0.271			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 11.2.5
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=44)	42	2.35 (1.459)	0.225	2.14 (1.394)	0.215	-0.20 (0.690)	0.106			0.044
Interferon beta-1a i.m. (N=52)	44	2.61 (1.234)	0.186	2.69 (1.308)	0.197	0.08 (0.731)	0.110			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.2.6
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=116)	107	2.53 (1.323)	0.128	2.48 (1.422)	0.137	-0.05 (0.795)	0.077	0.614		
Interferon beta-1a i.m. (N=97)	80	2.28 (1.174)	0.131	2.29 (1.272)	0.142	0.02 (0.858)	0.096			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 11.2.7
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=129)	119	1.97 (0.979)	0.090	1.91 (1.077)	0.099	-0.06 (0.737)	0.068	0.191		
Interferon beta-1a i.m. (N=125)	105	2.05 (0.951)	0.093	2.12 (1.073)	0.105	0.07 (0.806)	0.079			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.2.8
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=31)	30	4.48 (0.565)	0.103	4.25 (1.006)	0.184	-0.23 (0.878)	0.160	0.396		
Interferon beta-1a i.m. (N=24)	19	4.29 (0.419)	0.096	4.18 (0.989)	0.227	-0.11 (0.859)	0.197			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.3.1
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=19)	14	2.46 (1.293)	0.346	2.21 (1.637)	0.438	-0.25 (0.849)	0.227	0.243		
Placebo (N=19)	12	2.58 (1.145)	0.330	2.75 (1.699)	0.490	0.17 (0.778)	0.225			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 11.3.2
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=41)	37	2.26 (1.065)	0.175	2.31 (1.459)	0.240	0.05 (1.039)	0.171	0.761		
Placebo (N=35)	22	3.32 (1.715)	0.366	3.11 (1.661)	0.354	-0.20 (1.031)	0.220			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.3.3
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=51)	45	2.07 (0.939)	0.140	2.08 (1.373)	0.205	0.01 (0.962)	0.143	0.472		
Placebo (N=34)	20	1.95 (0.916)	0.205	1.95 (1.037)	0.232	0.00 (0.889)	0.199			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.3.4
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=9)	6	4.17 (0.408)	0.167	3.83 (1.571)	0.641	-0.33 (1.252)	0.511	0.494		
Placebo (N=20)	14	4.64 (0.602)	0.161	4.46 (1.184)	0.317	-0.18 (1.067)	0.285			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.3.5
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=44)	43	2.48 (1.460)	0.223	2.28 (1.415)	0.216	-0.20 (0.674)	0.103	0.040		
Interferon beta-1a i.m. (N=45)	38	2.64 (1.219)	0.198	2.72 (1.250)	0.203	0.08 (0.740)	0.120			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL

Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.3.6
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=122)	111	2.42 (1.290)	0.122	2.35 (1.339)	0.127	-0.07 (0.765)	0.073	0.450		
Interferon beta-1a i.m. (N=93)	76	2.32 (1.177)	0.135	2.34 (1.327)	0.152	0.03 (0.890)	0.102			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL

Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.3.7
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=134)	123	1.93 (0.937)	0.085	1.87 (1.002)	0.090	-0.07 (0.718)	0.065	0.136		
Interferon beta-1a i.m. (N=115)	96	2.08 (0.947)	0.097	2.15 (1.098)	0.112	0.07 (0.833)	0.085			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for gender.

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL

Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.3.8
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=32)	31	4.45 (0.522)	0.094	4.18 (0.945)	0.170	-0.27 (0.815)	0.146	0.368		
Interferon beta-1a i.m. (N=23)	18	4.28 (0.428)	0.101	4.17 (1.015)	0.239	-0.11 (0.884)	0.208			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for gender.

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL

Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 11.4.1
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=15)	12	1.92 (1.203)	0.347	2.04 (1.630)	0.471	0.13 (1.025)	0.296	0.327		
Placebo (N=13)	10	1.90 (1.350)	0.427	1.25 (1.379)	0.436	-0.65 (0.973)	0.308			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.4.2
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=33)	29	2.33 (1.219)	0.226	2.22 (1.544)	0.287	-0.10 (0.920)	0.171	0.224		
Placebo (N=24)	16	2.25 (0.931)	0.233	2.59 (1.369)	0.342	0.34 (0.978)	0.245			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 11.4.3
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=41)	36	1.90 (0.924)	0.154	1.86 (1.323)	0.221	-0.04 (0.981)	0.164	0.759		
Placebo (N=33)	23	1.85 (0.845)	0.176	1.76 (1.224)	0.255	-0.09 (1.041)	0.217			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.4.4
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=7)	5	4.40 (0.652)	0.292	4.40 (1.294)	0.579	0.00 (0.707)	0.316	0.635		
Placebo (N=4)	3	4.17 (0.289)	0.167	4.50 (1.323)	0.764	0.33 (1.528)	0.882			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.4.5
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=7)	6	1.58 (0.861)	0.352	1.42 (0.801)	0.327	-0.17 (0.258)	0.105	0.542		
Interferon beta-1a i.m. (N=7)	7	2.36 (1.180)	0.446	2.07 (1.367)	0.517	-0.29 (0.756)	0.286			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL

Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 11.4.6
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=20)	18	1.89 (0.948)	0.223	1.86 (0.904)	0.213	-0.03 (0.320)	0.075	0.005		
Interferon beta-1a i.m. (N=23)	20	1.93 (1.600)	0.358	2.25 (1.446)	0.323	0.33 (0.654)	0.146			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for baseline EDSS (<=3.5, >3.5).

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL

Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 11.4.7
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=26)	23	1.72 (0.809)	0.169	1.70 (0.863)	0.180	-0.02 (0.237)	0.049	0.087		
Interferon beta-1a i.m. (N=25)	22	1.52 (1.085)	0.231	1.73 (1.009)	0.215	0.20 (0.766)	0.163			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for gender.

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL

Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 11.4.8
 Change from baseline in EDSS, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]			
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c	
Month 12											
FTY720 0.5mg (N=1)	1	4.00 (N/A)	N/A	3.00 (N/A)	N/A	-1.00 (N/A)	N/A				0.012
Interferon beta-1a i.m. (N=5)	5	4.30 (0.671)	0.300	4.30 (0.837)	0.374	0.00 (0.500)	0.224				

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for gender.

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.1.1
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=129)	104	0.14 (0.489)	0.048	0.17 (0.608)	0.060	0.03 (0.424)	0.042	0.908		
Placebo (N=120)	87	0.03 (0.588)	0.063	0.09 (0.566)	0.061	0.05 (0.310)	0.033			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (≤ 3.5 , > 3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.1.2
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=296)	257	0.06 (0.590)	0.037	0.09 (0.734)	0.046	0.03 (0.382)	0.024	0.001		
Placebo (N=298)	229	0.01 (0.634)	0.042	-0.09 (0.986)	0.065	-0.10 (0.637)	0.042			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.1.3
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=363)	311	0.17 (0.496)	0.028	0.22 (0.594)	0.034	0.05 (0.372)	0.021	0.038		
Placebo (N=346)	272	0.12 (0.541)	0.033	0.12 (0.746)	0.045	0.00 (0.530)	0.032			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.1.4
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=62)	50	-0.47 (0.645)	0.091	-0.57 (0.901)	0.127	-0.10 (0.496)	0.070			
Placebo (N=72)	44	-0.63 (0.697)	0.105	-1.05 (1.058)	0.159	-0.42 (0.670)	0.101			0.035

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 12.1.5
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=148)	135	0.08 (0.532)	0.046	0.09 (0.557)	0.048	0.01 (0.358)	0.031	0.010		
Interferon beta-1a i.m. (N=139)	125	-0.02 (0.640)	0.057	-0.12 (0.722)	0.065	-0.10 (0.670)	0.060			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.1.6
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=281)	248	0.03 (0.641)	0.041	0.08 (0.591)	0.038	0.05 (0.448)	0.028	0.169		
Interferon beta-1a i.m. (N=292)	241	0.08 (0.556)	0.036	0.09 (0.526)	0.034	0.01 (0.334)	0.022			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.1.7
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=362)	321	0.15 (0.477)	0.027	0.20 (0.466)	0.026	0.04 (0.319)	0.018	0.023		
Interferon beta-1a i.m. (N=371)	316	0.11 (0.521)	0.029	0.11 (0.524)	0.029	-0.01 (0.371)	0.021			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 08SEP11 9:09

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 12.1.8
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=67)	62	-0.52 (0.848)	0.108	-0.52 (0.720)	0.091	-0.00 (0.749)	0.095	0.196		
Interferon beta-1a i.m. (N=60)	50	-0.39 (0.773)	0.109	-0.53 (0.788)	0.111	-0.15 (0.896)	0.127			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 12.2.1
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=19)	13	0.18 (0.614)	0.170	0.10 (0.642)	0.178	-0.08 (0.503)	0.139	0.296		
Placebo (N=14)	7	0.03 (0.537)	0.203	0.09 (0.633)	0.239	0.06 (0.276)	0.104			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 12.2.2
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=38)	33	-0.03 (0.572)	0.099	-0.02 (0.763)	0.133	0.02 (0.310)	0.054	0.391		
Placebo (N=38)	26	-0.16 (0.776)	0.152	-0.47 (1.121)	0.220	-0.32 (0.572)	0.112			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.2.3
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=49)	41	0.11 (0.525)	0.082	0.11 (0.598)	0.093	0.00 (0.301)	0.047	0.726		
Placebo (N=31)	21	0.18 (0.463)	0.101	0.17 (0.651)	0.142	-0.01 (0.395)	0.086			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.2.4
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=8)	5	-0.65 (0.669)	0.299	-0.78 (1.206)	0.539	-0.13 (0.788)	0.353	0.259		
Placebo (N=21)	12	-0.64 (0.823)	0.238	-1.27 (1.004)	0.290	-0.63 (0.553)	0.160			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 12.2.5
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=44)	42	0.02 (0.525)	0.081	0.00 (0.517)	0.080	-0.02 (0.314)	0.048	0.063		
Interferon beta-1a i.m. (N=52)	44	-0.04 (0.629)	0.095	-0.35 (0.824)	0.124	-0.31 (0.968)	0.146			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 12.2.6
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=116)	103	-0.10 (0.588)	0.058	-0.06 (0.663)	0.065	0.04 (0.339)	0.033	0.031		
Interferon beta-1a i.m. (N=97)	80	0.06 (0.551)	0.062	0.03 (0.514)	0.057	-0.03 (0.281)	0.031			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.2.7
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=129)	116	0.07 (0.477)	0.044	0.08 (0.538)	0.050	0.02 (0.324)	0.030	0.040		
Interferon beta-1a i.m. (N=125)	105	0.10 (0.456)	0.044	0.02 (0.526)	0.051	-0.08 (0.403)	0.039			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 12.2.8
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=31)	29	-0.57 (0.640)	0.119	-0.55 (0.689)	0.128	0.02 (0.370)	0.069	0.023		
Interferon beta-1a i.m. (N=24)	19	-0.37 (0.948)	0.218	-0.78 (0.915)	0.210	-0.41 (1.292)	0.296			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.3.1
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=19)	13	0.20 (0.612)	0.170	0.09 (0.632)	0.175	-0.11 (0.469)	0.130	0.104		
Placebo (N=19)	10	0.17 (0.623)	0.197	0.23 (0.640)	0.202	0.07 (0.245)	0.077			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

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

Novartis

FTY720 HTA

HTA Table 12.3.2
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=41)	36	0.05 (0.587)	0.098	0.02 (0.769)	0.128	-0.03 (0.380)	0.063	0.081		
Placebo (N=35)	21	-0.24 (0.813)	0.177	-0.67 (1.128)	0.246	-0.44 (0.553)	0.121			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.3.3
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=51)	43	0.16 (0.527)	0.080	0.11 (0.610)	0.093	-0.05 (0.353)	0.054	0.917		
Placebo (N=34)	20	0.25 (0.506)	0.113	0.21 (0.660)	0.148	-0.04 (0.364)	0.081			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.3.4
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=9)	6	-0.43 (0.815)	0.333	-0.52 (1.249)	0.510	-0.10 (0.709)	0.290	0.105		
Placebo (N=20)	11	-0.76 (0.762)	0.230	-1.46 (0.820)	0.247	-0.70 (0.520)	0.157			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.3.5
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=44)	43	0.08 (0.500)	0.076	0.01 (0.633)	0.096	-0.06 (0.442)	0.067	0.034		
Interferon beta-1a i.m. (N=45)	38	-0.14 (0.498)	0.081	-0.35 (0.758)	0.123	-0.21 (0.634)	0.103			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

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

Novartis

FTY720 HTA

HTA Table 12.3.6
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=122)	108	-0.06 (0.601)	0.058	-0.05 (0.671)	0.065	0.01 (0.342)	0.033			
Interferon beta-1a i.m. (N=93)	76	0.07 (0.543)	0.062	0.03 (0.516)	0.059	-0.03 (0.282)	0.032			0.049

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 12.3.7
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=134)	121	0.10 (0.479)	0.044	0.11 (0.541)	0.049	0.01 (0.323)	0.029			0.031
Interferon beta-1a i.m. (N=115)	96	0.10 (0.447)	0.046	0.01 (0.532)	0.054	-0.08 (0.415)	0.042			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.3.8
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=32)	30	-0.54 (0.651)	0.119	-0.60 (0.784)	0.143	-0.06 (0.535)	0.098	0.075		
Interferon beta-1a i.m. (N=23)	18	-0.53 (0.660)	0.156	-0.67 (0.803)	0.189	-0.14 (0.552)	0.130			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 12.4.1
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=15)	12	0.16 (0.187)	0.054	0.15 (0.743)	0.215	-0.01 (0.740)	0.214	0.530		
Placebo (N=13)	10	0.19 (0.377)	0.119	0.23 (0.384)	0.121	0.03 (0.346)	0.109			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 12.4.2
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=33)	26	0.10 (0.482)	0.095	0.21 (0.576)	0.113	0.11 (0.376)	0.074	0.160		
Placebo (N=24)	15	-0.05 (0.772)	0.199	-0.28 (1.289)	0.333	-0.23 (0.813)	0.210			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.4.3
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=41)	34	0.16 (0.400)	0.069	0.24 (0.607)	0.104	0.08 (0.530)	0.091	0.298		
Placebo (N=33)	22	0.15 (0.558)	0.119	0.14 (0.530)	0.113	-0.01 (0.388)	0.083			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 12.4.4
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=7)	4	-0.21 (0.389)	0.195	-0.27 (0.659)	0.329	-0.06 (0.339)	0.169	0.455		
Placebo (N=4)	3	-0.67 (0.904)	0.522	-1.68 (2.391)	1.381	-1.01 (1.574)	0.909			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 12.4.5
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=7)	6	0.20 (0.351)	0.143	0.19 (0.432)	0.176	-0.00 (0.193)	0.079	0.650		
Interferon beta-1a i.m. (N=7)	7	0.03 (0.537)	0.203	-0.11 (0.546)	0.206	-0.13 (0.293)	0.111			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 12.4.6
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=20)	15	0.17 (0.451)	0.117	0.20 (0.380)	0.098	0.03 (0.140)	0.036	0.354		
Interferon beta-1a i.m. (N=23)	19	0.00 (0.647)	0.148	0.09 (0.584)	0.134	0.09 (0.306)	0.070			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 12.4.7
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=26)	20	0.22 (0.376)	0.084	0.24 (0.342)	0.076	0.02 (0.156)	0.035	0.666		
Interferon beta-1a i.m. (N=25)	21	0.12 (0.542)	0.118	0.13 (0.577)	0.126	0.01 (0.269)	0.059			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 12.4.8
 Change from baseline in MSFC z-score, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=1)	1	-0.69 (N/A)	N/A	-0.64 (N/A)	N/A	0.05 (N/A)	N/A	0.787		
Interferon beta-1a i.m. (N=5)	5	-0.46 (0.709)	0.317	-0.34 (0.386)	0.173	0.12 (0.490)	0.219			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline MSFC z-score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.1.1
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=129)	106	0.83 (0.150)	0.015	0.82 (0.181)	0.018	-0.01 (0.159)	0.015	0.121		
Placebo (N=120)	91	0.82 (0.170)	0.018	0.79 (0.206)	0.022	-0.03 (0.174)	0.018			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.1.2
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=296)	260	0.79 (0.191)	0.012	0.78 (0.200)	0.012	-0.00 (0.167)	0.010	0.631		
Placebo (N=298)	235	0.73 (0.199)	0.013	0.76 (0.207)	0.014	0.03 (0.197)	0.013			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.1.3
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=363)	315	0.82 (0.166)	0.009	0.81 (0.189)	0.011	-0.01 (0.159)	0.009	0.893		
Placebo (N=346)	274	0.79 (0.171)	0.010	0.80 (0.193)	0.012	0.01 (0.187)	0.011			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 13.1.4
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=62)	51	0.64 (0.192)	0.027	0.67 (0.191)	0.027	0.03 (0.191)	0.027	0.104		
Placebo (N=72)	52	0.61 (0.242)	0.034	0.64 (0.225)	0.031	0.03 (0.217)	0.030			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.1.5
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=148)	128	0.82 (0.204)	0.018	0.82 (0.210)	0.019	0.00 (0.161)	0.014	0.183		
Interferon beta-1a i.m. (N=139)	115	0.82 (0.153)	0.014	0.81 (0.212)	0.020	-0.01 (0.189)	0.018			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 13.1.6
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=281)	246	0.78 (0.190)	0.012	0.80 (0.229)	0.015	0.01 (0.187)	0.012	0.384		
Interferon beta-1a i.m. (N=292)	242	0.81 (0.175)	0.011	0.81 (0.194)	0.012	0.00 (0.181)	0.012			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.1.7
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=362)	316	0.82 (0.174)	0.010	0.83 (0.199)	0.011	0.01 (0.170)	0.010	0.370		
Interferon beta-1a i.m. (N=371)	314	0.83 (0.151)	0.009	0.84 (0.184)	0.010	0.00 (0.176)	0.010			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 08SEP11 9:09

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.1.8
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=67)	58	0.65 (0.239)	0.031	0.65 (0.277)	0.036	-0.00 (0.223)	0.029	0.086		
Interferon beta-1a i.m. (N=60)	43	0.68 (0.218)	0.033	0.63 (0.217)	0.033	-0.05 (0.226)	0.034			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL

Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.2.1
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=19)	13	0.78 (0.155)	0.043	0.72 (0.200)	0.056	-0.05 (0.182)	0.051	0.404		
Placebo (N=14)	9	0.82 (0.172)	0.057	0.81 (0.152)	0.051	-0.01 (0.159)	0.053			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 13.2.2
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=38)	33	0.73 (0.193)	0.034	0.74 (0.245)	0.043	0.01 (0.227)	0.040	0.182		
Placebo (N=38)	26	0.75 (0.153)	0.030	0.82 (0.123)	0.024	0.07 (0.186)	0.037			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.2.3
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=49)	41	0.75 (0.182)	0.028	0.75 (0.241)	0.038	0.00 (0.223)	0.035	0.143		
Placebo (N=31)	21	0.82 (0.168)	0.037	0.88 (0.115)	0.025	0.06 (0.214)	0.047			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.2.4
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=8)	5	0.69 (0.210)	0.094	0.64 (0.106)	0.048	-0.05 (0.154)	0.069	0.289		
Placebo (N=21)	14	0.70 (0.118)	0.032	0.73 (0.091)	0.024	0.02 (0.120)	0.032			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 13.2.5
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=44)	38	0.83 (0.182)	0.030	0.82 (0.170)	0.028	-0.01 (0.161)	0.026	0.717		
Interferon beta-1a i.m. (N=52)	38	0.83 (0.159)	0.026	0.83 (0.171)	0.028	0.00 (0.173)	0.028			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.2.6
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=116)	100	0.77 (0.189)	0.019	0.78 (0.253)	0.025	0.01 (0.188)	0.019	0.870		
Interferon beta-1a i.m. (N=97)	80	0.78 (0.189)	0.021	0.82 (0.168)	0.019	0.04 (0.165)	0.018			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.2.7
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=129)	110	0.81 (0.193)	0.018	0.82 (0.224)	0.021	0.02 (0.160)	0.015	0.828		
Interferon beta-1a i.m. (N=125)	103	0.82 (0.161)	0.016	0.85 (0.156)	0.015	0.03 (0.168)	0.017			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.2.8
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=31)	28	0.71 (0.142)	0.027	0.67 (0.235)	0.044	-0.03 (0.243)	0.046			0.451
Interferon beta-1a i.m. (N=24)	15	0.64 (0.230)	0.059	0.66 (0.162)	0.042	0.02 (0.175)	0.045			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 13.3.1
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=19)	13	0.80 (0.166)	0.046	0.74 (0.214)	0.059	-0.06 (0.181)	0.050	0.468		
Placebo (N=19)	12	0.84 (0.169)	0.049	0.81 (0.146)	0.042	-0.02 (0.101)	0.029			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.3.2
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=41)	36	0.72 (0.227)	0.038	0.73 (0.261)	0.044	0.01 (0.228)	0.038	0.818		
Placebo (N=35)	20	0.78 (0.124)	0.028	0.80 (0.107)	0.024	0.01 (0.123)	0.027			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Novartis

FTY720 HTA

HTA Table 13.3.3
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=51)	43	0.75 (0.217)	0.033	0.74 (0.259)	0.040	-0.01 (0.224)	0.034	0.513		
Placebo (N=34)	19	0.87 (0.122)	0.028	0.86 (0.108)	0.025	-0.01 (0.115)	0.026			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.3.4
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=9)	6	0.69 (0.188)	0.077	0.68 (0.137)	0.056	-0.01 (0.170)	0.069	0.828		
Placebo (N=20)	13	0.71 (0.121)	0.033	0.72 (0.093)	0.026	0.01 (0.118)	0.033			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 13.3.5
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=44)	39	0.82 (0.164)	0.026	0.81 (0.158)	0.025	-0.01 (0.145)	0.023	0.499		
Interferon beta-1a i.m. (N=45)	31	0.82 (0.157)	0.028	0.82 (0.180)	0.032	-0.01 (0.188)	0.034			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 13.3.6
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=122)	108	0.77 (0.203)	0.020	0.79 (0.254)	0.024	0.02 (0.183)	0.018	0.634		
Interferon beta-1a i.m. (N=93)	76	0.77 (0.191)	0.022	0.81 (0.170)	0.020	0.04 (0.173)	0.020			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.3.7
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=134)	118	0.80 (0.192)	0.018	0.83 (0.211)	0.019	0.02 (0.154)	0.014	0.553		
Interferon beta-1a i.m. (N=115)	93	0.81 (0.164)	0.017	0.84 (0.161)	0.017	0.03 (0.178)	0.018			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.3.8
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=32)	29	0.69 (0.178)	0.033	0.65 (0.265)	0.049	-0.03 (0.234)	0.044	0.589		
Interferon beta-1a i.m. (N=23)	14	0.65 (0.238)	0.064	0.66 (0.168)	0.045	0.01 (0.180)	0.048			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.4.1
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=15)	12	0.85 (0.129)	0.037	0.90 (0.125)	0.036	0.06 (0.103)	0.030	0.852		
Placebo (N=13)	9	0.83 (0.115)	0.038	0.88 (0.150)	0.050	0.05 (0.120)	0.040			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.4.2
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=33)	28	0.81 (0.170)	0.032	0.83 (0.206)	0.039	0.02 (0.186)	0.035	0.002		
Placebo (N=24)	15	0.78 (0.203)	0.052	0.63 (0.314)	0.081	-0.15 (0.206)	0.053			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 13.4.3
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=41)	35	0.85 (0.130)	0.022	0.87 (0.184)	0.031	0.01 (0.166)	0.028	0.058		
Placebo (N=33)	21	0.82 (0.125)	0.027	0.76 (0.251)	0.055	-0.06 (0.212)	0.046			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 13.4.4
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=7)	5	0.59 (0.145)	0.065	0.72 (0.174)	0.078	0.14 (0.122)	0.054	<0.001		
Placebo (N=4)	3	0.62 (0.373)	0.215	0.47 (0.475)	0.274	-0.15 (0.106)	0.061			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.4.5
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=7)	6	0.90 (0.161)	0.066	0.89 (0.195)	0.080	-0.01 (0.123)	0.050	0.437		
Interferon beta-1a i.m. (N=7)	6	0.86 (0.160)	0.065	0.86 (0.190)	0.078	0.00 (0.189)	0.077			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.4.6
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=20)	17	0.81 (0.148)	0.036	0.87 (0.112)	0.027	0.06 (0.105)	0.026	0.029		
Interferon beta-1a i.m. (N=23)	17	0.86 (0.168)	0.041	0.83 (0.151)	0.037	-0.03 (0.127)	0.031			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.4.7
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=26)	22	0.85 (0.149)	0.032	0.88 (0.133)	0.028	0.04 (0.112)	0.024	0.233		
Interferon beta-1a i.m. (N=25)	20	0.89 (0.151)	0.034	0.87 (0.139)	0.031	-0.02 (0.139)	0.031			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 13.4.8
 Change from baseline in utility scores converted from EQ-5D, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=1)	1	0.62 (N/A)	N/A	0.74 (N/A)	N/A	0.12 (N/A)	N/A	0.298		
Interferon beta-1a i.m. (N=5)	3	0.68 (0.141)	0.081	0.61 (0.089)	0.051	-0.06 (0.184)	0.106			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline utility score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 14.1.1
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=129)	105	78.54 (14.41)	1.406	77.30 (15.69)	1.532	-1.24 (13.71)	1.338	0.149		
Placebo (N=120)	92	77.73 (14.47)	1.508	74.33 (16.80)	1.751	-3.40 (15.10)	1.574			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.1.2
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=296)	260	76.62 (17.08)	1.059	76.80 (17.01)	1.055	0.18 (16.32)	1.012	0.832		
Placebo (N=298)	234	74.56 (16.83)	1.100	74.45 (18.26)	1.194	-0.11 (17.66)	1.154			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.1.3
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=363)	314	79.67 (15.03)	0.848	78.61 (15.75)	0.889	-1.06 (15.23)	0.860	0.992		
Placebo (N=346)	275	77.43 (14.92)	0.900	76.75 (16.62)	1.002	-0.68 (16.50)	0.995			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.1.4
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=62)	51	61.80 (15.98)	2.238	66.69 (18.28)	2.559	4.88 (17.01)	2.381	0.049		
Placebo (N=72)	51	64.82 (18.89)	2.645	61.82 (19.03)	2.665	-3.00 (19.64)	2.750			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 14.1.5
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=148)	130	76.66 (16.69)	1.464	79.04 (14.73)	1.292	2.38 (13.40)	1.175	0.558		
Interferon beta-1a i.m. (N=139)	115	77.78 (16.22)	1.513	78.85 (16.28)	1.519	1.07 (13.16)	1.227			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 14.1.6
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=281)	245	78.13 (16.16)	1.032	79.76 (16.66)	1.064	1.63 (15.51)	0.991	0.335		
Interferon beta-1a i.m. (N=292)	239	78.63 (16.13)	1.044	79.05 (16.43)	1.063	0.42 (15.20)	0.983			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.1.7
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=362)	317	79.60 (15.19)	0.853	81.33 (15.08)	0.847	1.74 (14.78)	0.830			0.713
Interferon beta-1a i.m. (N=371)	312	80.49 (14.58)	0.825	81.62 (14.14)	0.800	1.13 (13.50)	0.764			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL

Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.1.8
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=67)	58	66.81 (18.21)	2.391	69.55 (17.30)	2.272	2.74 (15.02)	1.972	0.034		
Interferon beta-1a i.m. (N=60)	42	62.50 (18.45)	2.847	59.43 (18.54)	2.861	-3.07 (20.64)	3.185			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 14.2.1
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=19)	13	74.69 (12.90)	3.578	71.15 (17.67)	4.900	-3.54 (16.32)	4.527	0.447		
Placebo (N=14)	9	85.67 (12.84)	4.278	81.78 (13.68)	4.561	-3.89 (13.19)	4.395			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.2.2
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=38)	33	72.73 (14.61)	2.544	75.55 (16.89)	2.941	2.82 (16.77)	2.919	0.133		
Placebo (N=38)	25	76.20 (18.22)	3.644	80.84 (11.70)	2.339	4.64 (18.54)	3.708			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.2.3
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=49)	41	74.29 (13.35)	2.085	75.07 (17.47)	2.728	0.78 (17.17)	2.682	0.119		
Placebo (N=31)	21	81.95 (16.85)	3.678	86.38 (10.29)	2.245	4.43 (15.10)	3.295			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.2.4
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=8)	5	65.00 (18.40)	8.228	68.00 (12.55)	5.612	3.00 (13.73)	6.140	0.528		
Placebo (N=21)	13	73.46 (17.37)	4.817	72.54 (9.709)	2.693	-0.92 (21.08)	5.846			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.2.5
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=44)	40	76.30 (16.29)	2.576	79.18 (13.79)	2.181	2.88 (12.83)	2.028	0.406		
Interferon beta-1a i.m. (N=52)	38	77.21 (14.48)	2.349	76.92 (13.56)	2.199	-0.29 (10.58)	1.716			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.2.6
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=116)	99	76.83 (15.62)	1.569	80.18 (16.42)	1.650	3.35 (14.07)	1.414	0.012		
Interferon beta-1a i.m. (N=97)	80	76.01 (17.64)	1.972	74.80 (17.88)	1.999	-1.21 (16.22)	1.813			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.2.7
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=129)	111	79.19 (14.00)	1.329	82.26 (14.48)	1.375	3.07 (13.60)	1.291	0.061		
Interferon beta-1a i.m. (N=125)	103	78.72 (15.37)	1.514	78.21 (14.67)	1.445	-0.50 (14.61)	1.439			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.2.8
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=31)	28	66.71 (18.46)	3.490	70.50 (16.88)	3.191	3.79 (14.24)	2.691	0.015		
Interferon beta-1a i.m. (N=24)	15	60.47 (16.72)	4.318	56.73 (17.31)	4.469	-3.73 (14.72)	3.802			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.3.1
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=19)	13	75.92 (14.06)	3.898	72.92 (18.87)	5.234	-3.00 (16.31)	4.525	0.892		
Placebo (N=19)	12	84.00 (17.07)	4.927	81.67 (13.65)	3.940	-2.33 (13.98)	4.036			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.3.2
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=41)	36	72.14 (18.86)	3.143	75.00 (16.22)	2.703	2.86 (20.51)	3.419	0.442		
Placebo (N=35)	19	75.79 (19.67)	4.514	78.89 (11.74)	2.693	3.11 (20.39)	4.677			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.3.3
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=51)	43	74.63 (17.35)	2.646	75.30 (17.34)	2.645	0.67 (20.19)	3.078	0.358		
Placebo (N=34)	19	82.00 (19.26)	4.418	84.53 (11.64)	2.671	2.53 (15.71)	3.604			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.3.4
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=9)	6	62.50 (17.56)	7.168	68.33 (11.25)	4.595	5.83 (14.11)	5.759	0.489		
Placebo (N=20)	12	74.17 (17.94)	5.180	72.75 (10.11)	2.918	-1.42 (21.94)	6.332			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.3.5
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=44)	41	75.02 (15.75)	2.459	76.95 (14.17)	2.214	1.93 (12.15)	1.898	0.942		
Interferon beta-1a i.m. (N=45)	31	77.26 (15.33)	2.754	77.26 (14.00)	2.514	0.00 (11.31)	2.031			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.3.6
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=122)	107	77.83 (14.85)	1.435	81.18 (15.59)	1.507	3.35 (13.71)	1.325	0.002		
Interferon beta-1a i.m. (N=93)	76	76.17 (17.16)	1.968	74.20 (18.04)	2.070	-1.97 (16.17)	1.855			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.3.7
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=134)	119	79.48 (13.28)	1.217	82.30 (14.27)	1.308	2.82 (13.09)	1.200	0.031		
Interferon beta-1a i.m. (N=115)	93	79.05 (15.03)	1.558	77.92 (14.97)	1.553	-1.13 (14.93)	1.548			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.3.8
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=32)	29	67.10 (18.06)	3.354	70.59 (15.91)	2.955	3.48 (14.21)	2.639	0.052		
Interferon beta-1a i.m. (N=23)	14	59.43 (16.85)	4.502	56.21 (17.84)	4.768	-3.21 (15.14)	4.045			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.4.1
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=15)	12	81.67 (7.177)	2.072	86.50 (9.691)	2.797	4.83 (9.340)	2.696	0.431		
Placebo (N=13)	10	79.50 (12.79)	4.045	83.00 (9.463)	2.993	3.50 (12.60)	3.984			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.4.2
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=33)	28	77.68 (15.41)	2.912	78.57 (17.41)	3.290	0.89 (17.69)	3.344	0.061		
Placebo (N=24)	15	77.40 (19.47)	5.028	70.13 (24.37)	6.292	-7.27 (27.10)	6.997			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.4.3
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=41)	35	81.00 (10.54)	1.781	82.37 (15.06)	2.545	1.37 (16.16)	2.731	0.046		
Placebo (N=33)	22	80.45 (16.28)	3.471	77.36 (15.99)	3.408	-3.09 (22.07)	4.705			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Novartis

FTY720 HTA

HTA Table 14.4.4
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 24										
FTY720 0.5mg (N=7)	5	64.00 (22.75)	10.17	71.00 (19.17)	8.573	7.00 (11.51)	5.148	0.894		
Placebo (N=4)	3	62.00 (13.11)	7.572	60.00 (44.44)	25.66	-2.00 (32.97)	19.04			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.
 a= FTY720 1.25mg vs. Placebo, b=FTY720 0.5mg vs. Placebo, c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.4.5
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Males

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=7)	6	79.67 (22.87)	9.337	82.50 (19.71)	8.045	2.83 (7.885)	3.219	0.527		
Interferon beta-1a i.m. (N=7)	6	83.17 (15.37)	6.274	76.83 (31.09)	12.69	-6.33 (25.10)	10.25			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 14.4.6
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Females

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=20)	17	79.94 (21.78)	5.281	82.71 (15.95)	3.869	2.76 (26.46)	6.417	0.236		
Interferon beta-1a i.m. (N=23)	16	81.31 (13.18)	3.296	83.69 (11.91)	2.977	2.38 (6.801)	1.700			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and baseline EDSS (<=3.5, >3.5).
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 14.4.7
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS 0.0-3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=26)	22	79.86 (22.04)	4.699	84.14 (15.27)	3.256	4.27 (22.24)	4.742	0.461		
Interferon beta-1a i.m. (N=25)	19	84.21 (11.88)	2.725	86.05 (11.10)	2.546	1.84 (6.167)	1.415			

For each visit, only patients with non-missing baseline and post-baseline values are included.

[1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.

a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.

SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 14.4.8
 Change from baseline in EQ-5D Visual Analog Scale Scores, by visit and treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS >3.5

Treatment	n	Baseline		Post-baseline		Change from baseline		P-value [1]		
		Mean (SD)	SEM	Mean (SD)	SEM	Mean (SD)	SEM	a	b	c
Month 12										
FTY720 0.5mg (N=1)	1	80.00 (N/A)	N/A	50.00 (N/A)	N/A	-30.00 (N/A)	N/A	0.480		
Interferon beta-1a i.m. (N=5)	3	66.67 (15.28)	8.819	55.00 (35.00)	20.21	-11.67 (38.84)	22.42			

For each visit, only patients with non-missing baseline and post-baseline values are included.
 [1] Pairwise comparison, ANCOVA on ranks adjusted for baseline VAS score and gender.
 a= FTY720 1.25mg vs. Interferon beta-1a i.m., b=FTY720 0.5mg vs. Interferon beta-1a i.m., c=FTY720 1.25mg vs. FTY720 0.5mg.
 SD = Standard deviation. SEM = Standard error of the mean.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 11 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:03

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.1.1
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population
 Males

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	134	129	120
Total number of relapses	49	43	117
Severity, n(%)*			
Mild	15 (30.6)	16 (37.2)	33 (28.2)
Moderate	25 (51.0)	27 (62.8)	74 (63.2)
Severe	9 (18.4)	0 (0.0)	10 (8.5)
Steroid use, n(%)*	38 (77.6)	33 (76.7)	100 (85.5)
Hospitalisations, n(%)*	15 (30.6)	13 (30.2)	57 (48.7)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.1.2
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population
 Females

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	295	296	298
Total number of relapses	99	129	242
Severity, n(%)*			
Mild	28 (28.3)	44 (34.1)	86 (35.5)
Moderate	61 (61.6)	75 (58.1)	131 (54.1)
Severe	10 (10.1)	10 (7.8)	25 (10.3)
Steroid use, n(%)*	82 (82.8)	107 (82.9)	203 (83.9)
Hospitalisations, n(%)*	45 (45.5)	50 (38.8)	89 (36.8)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.1.3
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population
 Baseline EDSS 0.0-3.5

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	351	363	346
Total number of relapses	104	144	274
Severity, n(%)*			
Mild	28 (26.9)	52 (36.1)	95 (34.7)
Moderate	63 (60.6)	83 (57.6)	154 (56.2)
Severe	13 (12.5)	9 (6.3)	25 (9.1)
Steroid use, n(%)*	82 (78.8)	115 (79.9)	224 (81.8)
Hospitalisations, n(%)*	35 (33.7)	53 (36.8)	103 (37.6)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.1.4
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population
 Baseline EDSS >3.5

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	78	62	72
Total number of relapses	44	28	85
Severity, n(%)*			
Mild	15 (34.1)	8 (28.6)	24 (28.2)
Moderate	23 (52.3)	19 (67.9)	51 (60.0)
Severe	6 (13.6)	1 (3.6)	10 (11.8)
Steroid use, n(%)*	38 (86.4)	25 (89.3)	79 (92.9)
Hospitalisations, n(%)*	25 (56.8)	10 (35.7)	43 (50.6)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.1.5
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population
 Males

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	132	148	139
Total number of relapses	37	30	46
Severity, n(%)*			
Mild	14 (37.8)	11 (36.7)	10 (21.7)
Moderate	18 (48.6)	16 (53.3)	30 (65.2)
Severe	5 (13.5)	3 (10.0)	6 (13.0)
Steroid use, n(%)*	32 (86.5)	21 (70.0)	39 (84.8)
Hospitalisations, n(%)*	4 (10.8)	4 (13.3)	10 (21.7)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.1.6
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population
 Females

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	288	281	292
Total number of relapses	68	59	133
Severity, n(%)*			
Mild	20 (29.4)	20 (33.9)	34 (25.6)
Moderate	36 (52.9)	32 (54.2)	75 (56.4)
Severe	12 (17.6)	7 (11.9)	24 (18.0)
Steroid use, n(%)*	55 (80.9)	44 (74.6)	109 (82.0)
Hospitalisations, n(%)*	14 (20.6)	7 (11.9)	26 (19.5)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.1.7
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population
 Baseline EDSS 0.0-3.5

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	365	362	371
Total number of relapses	84	67	147
Severity, n(%)*			
Mild	27 (32.1)	23 (34.3)	29 (19.7)
Moderate	41 (48.8)	36 (53.7)	91 (61.9)
Severe	16 (19.0)	8 (11.9)	27 (18.4)
Steroid use, n(%)*	69 (82.1)	47 (70.1)	119 (81.0)
Hospitalisations, n(%)*	15 (17.9)	6 (9.0)	26 (17.7)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.1.8
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population
 Baseline EDSS >3.5

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	55	67	60
Total number of relapses	21	22	32
Severity, n(%)*			
Mild	7 (33.3)	8 (36.4)	15 (46.9)
Moderate	13 (61.9)	12 (54.5)	14 (43.8)
Severe	1 (4.8)	2 (9.1)	3 (9.4)
Steroid use, n(%)*	18 (85.7)	18 (81.8)	29 (90.6)
Hospitalisations, n(%)*	3 (14.3)	5 (22.7)	10 (31.3)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.2.1
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Males

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	14	19	14
Total number of relapses	6	7	17
Severity, n(%)*			
Mild	3 (50.0)	4 (57.1)	6 (35.3)
Moderate	2 (33.3)	3 (42.9)	8 (47.1)
Severe	1 (16.7)	0 (0.0)	3 (17.6)
Steroid use, n(%)*	5 (83.3)	4 (57.1)	14 (82.4)
Hospitalisations, n(%)*	2 (33.3)	2 (28.6)	8 (47.1)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.2.2

MS relapse (confirmed relapses only) characteristics by treatment

CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)

Females

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	31	38	38
Total number of relapses	10	18	37
Severity, n(%)*			
Mild	3 (30.0)	6 (33.3)	16 (43.2)
Moderate	5 (50.0)	11 (61.1)	17 (45.9)
Severe	2 (20.0)	1 (5.6)	4 (10.8)
Steroid use, n(%)*	8 (80.0)	12 (66.7)	31 (83.8)
Hospitalisations, n(%)*	2 (20.0)	6 (33.3)	9 (24.3)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.2.3

MS relapse (confirmed relapses only) characteristics by treatment

CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)

Baseline EDSS 0.0-3.5

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	32	49	31
Total number of relapses	8	24	22
Severity, n(%)*			
Mild	2 (25.0)	10 (41.7)	10 (45.5)
Moderate	5 (62.5)	13 (54.2)	8 (36.4)
Severe	1 (12.5)	1 (4.2)	4 (18.2)
Steroid use, n(%)*	7 (87.5)	15 (62.5)	13 (59.1)
Hospitalisations, n(%)*	4 (50.0)	8 (33.3)	5 (22.7)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL

Date/time of run: 05SEP11 18:19

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.2.4
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS >3.5

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	13	8	21
Total number of relapses	8	1	32
Severity, n(%)*			
Mild	4 (50.0)	0 (0.0)	12 (37.5)
Moderate	2 (25.0)	1 (100.0)	17 (53.1)
Severe	2 (25.0)	0 (0.0)	3 (9.4)
Steroid use, n(%)*	6 (75.0)	1 (100.0)	32 (100.0)
Hospitalisations, n(%)*	0 (0.0)	0 (0.0)	12 (37.5)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.2.5

MS relapse (confirmed relapses only) characteristics by treatment

FTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)

Males

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	37	44	52
Total number of relapses	16	10	24
Severity, n(%)*			
Mild	6 (37.5)	1 (10.0)	7 (29.2)
Moderate	7 (43.8)	7 (70.0)	13 (54.2)
Severe	3 (18.8)	2 (20.0)	4 (16.7)
Steroid use, n(%)*	15 (93.8)	9 (90.0)	22 (91.7)
Hospitalisations, n(%)*	1 (6.3)	0 (0.0)	4 (16.7)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.2.6
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Females

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	107	116	97
Total number of relapses	32	22	48
Severity, n(%)*			
Mild	10 (31.3)	5 (22.7)	11 (22.9)
Moderate	15 (46.9)	13 (59.1)	24 (50.0)
Severe	7 (21.9)	4 (18.2)	13 (27.1)
Steroid use, n(%)*	27 (84.4)	17 (77.3)	41 (85.4)
Hospitalisations, n(%)*	9 (28.1)	4 (18.2)	10 (20.8)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.2.7

MS relapse (confirmed relapses only) characteristics by treatment

CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)

Baseline EDSS 0.0-3.5

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	116	129	125
Total number of relapses	36	21	60
Severity, n(%)*			
Mild	12 (33.3)	4 (19.0)	12 (20.0)
Moderate	14 (38.9)	12 (57.1)	32 (53.3)
Severe	10 (27.8)	5 (23.8)	16 (26.7)
Steroid use, n(%)*	32 (88.9)	16 (76.2)	51 (85.0)
Hospitalisations, n(%)*	9 (25.0)	2 (9.5)	10 (16.7)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL

Date/time of run: 05SEP11 18:19

Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.2.8
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS >3.5

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	28	31	24
Total number of relapses	12	11	12
Severity, n(%)*			
Mild	4 (33.3)	2 (18.2)	6 (50.0)
Moderate	8 (66.7)	8 (72.7)	5 (41.7)
Severe	0 (0.0)	1 (9.1)	1 (8.3)
Steroid use, n(%)*	10 (83.3)	10 (90.9)	12 (100.0)
Hospitalisations, n(%)*	1 (8.3)	2 (18.2)	4 (33.3)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.3.1
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Males

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	15	19	19
Total number of relapses	11	5	17
Severity, n(%)*			
Mild	6 (54.5)	2 (40.0)	4 (23.5)
Moderate	4 (36.4)	3 (60.0)	10 (58.8)
Severe	1 (9.1)	0 (0.0)	3 (17.6)
Steroid use, n(%)*	10 (90.9)	4 (80.0)	16 (94.1)
Hospitalisations, n(%)*	5 (45.5)	2 (40.0)	8 (47.1)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.3.2
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Females

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	33	41	35
Total number of relapses	7	16	39
Severity, n(%)*			
Mild	2 (28.6)	7 (43.8)	16 (41.0)
Moderate	3 (42.9)	8 (50.0)	19 (48.7)
Severe	2 (28.6)	1 (6.3)	4 (10.3)
Steroid use, n(%)*	5 (71.4)	10 (62.5)	33 (84.6)
Hospitalisations, n(%)*	2 (28.6)	6 (37.5)	9 (23.1)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.3.3
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS 0.0-3.5

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	37	51	34
Total number of relapses	13	20	24
Severity, n(%)*			
Mild	5 (38.5)	9 (45.0)	8 (33.3)
Moderate	7 (53.8)	10 (50.0)	12 (50.0)
Severe	1 (7.7)	1 (5.0)	4 (16.7)
Steroid use, n(%)*	12 (92.3)	13 (65.0)	17 (70.8)
Hospitalisations, n(%)*	7 (53.8)	8 (40.0)	5 (20.8)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.3.4
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS >3.5

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	11	9	20
Total number of relapses	5	1	32
Severity, n(%)*			
Mild	3 (60.0)	0 (0.0)	12 (37.5)
Moderate	0 (0.0)	1 (100.0)	17 (53.1)
Severe	2 (40.0)	0 (0.0)	3 (9.4)
Steroid use, n(%)*	3 (60.0)	1 (100.0)	32 (100.0)
Hospitalisations, n(%)*	0 (0.0)	0 (0.0)	12 (37.5)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.3.5
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Males

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	44	44	45
Total number of relapses	20	12	19
Severity, n(%)*			
Mild	7 (35.0)	3 (25.0)	5 (26.3)
Moderate	9 (45.0)	8 (66.7)	10 (52.6)
Severe	4 (20.0)	1 (8.3)	4 (21.1)
Steroid use, n(%)*	19 (95.0)	10 (83.3)	17 (89.5)
Hospitalisations, n(%)*	1 (5.0)	0 (0.0)	2 (10.5)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.3.6
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Females

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	108	122	93
Total number of relapses	33	21	48
Severity, n(%)*			
Mild	8 (24.2)	5 (23.8)	10 (20.8)
Moderate	19 (57.6)	12 (57.1)	25 (52.1)
Severe	6 (18.2)	4 (19.0)	13 (27.1)
Steroid use, n(%)*	27 (81.8)	14 (66.7)	40 (83.3)
Hospitalisations, n(%)*	7 (21.2)	4 (19.0)	9 (18.8)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.3.7
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS 0.0-3.5

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	125	134	115
Total number of relapses	39	21	56
Severity, n(%)*			
Mild	11 (28.2)	5 (23.8)	10 (17.9)
Moderate	19 (48.7)	12 (57.1)	30 (53.6)
Severe	9 (23.1)	4 (19.0)	16 (28.6)
Steroid use, n(%)*	34 (87.2)	14 (66.7)	46 (82.1)
Hospitalisations, n(%)*	7 (17.9)	2 (9.5)	8 (14.3)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.3.8
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS >3.5

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	27	32	23
Total number of relapses	14	12	11
Severity, n(%)*			
Mild	4 (28.6)	3 (25.0)	5 (45.5)
Moderate	9 (64.3)	8 (66.7)	5 (45.5)
Severe	1 (7.1)	1 (8.3)	1 (9.1)
Steroid use, n(%)*	12 (85.7)	10 (83.3)	11 (100.0)
Hospitalisations, n(%)*	1 (7.1)	2 (16.7)	3 (27.3)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.4.1

MS relapse (confirmed relapses only) characteristics by treatment

CFTY720D2301 ITT population (Rapidly evolving: IFN naive, ≥ 2 relapses in year-1 and ≥ 1 Gd+ at baseline)

Males

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	16	15	13
Total number of relapses	11	7	14
Severity, n(%)*			
Mild	3 (27.3)	3 (42.9)	2 (14.3)
Moderate	4 (36.4)	4 (57.1)	12 (85.7)
Severe	4 (36.4)	0 (0.0)	0 (0.0)
Steroid use, n(%)*	8 (72.7)	7 (100.0)	13 (92.9)
Hospitalisations, n(%)*	2 (18.2)	3 (42.9)	8 (57.1)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.4.2
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Females

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	26	33	24
Total number of relapses	12	15	32
Severity, n(%)*			
Mild	6 (50.0)	1 (6.7)	8 (25.0)
Moderate	5 (41.7)	13 (86.7)	20 (62.5)
Severe	1 (8.3)	1 (6.7)	4 (12.5)
Steroid use, n(%)*	10 (83.3)	14 (93.3)	24 (75.0)
Hospitalisations, n(%)*	4 (33.3)	3 (20.0)	11 (34.4)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.4.3

MS relapse (confirmed relapses only) characteristics by treatment

CFTY720D2301 ITT population (Rapidly evolving: IFN naive, ≥ 2 relapses in year-1 and ≥ 1 Gd+ at baseline)

Baseline EDSS 0.0-3.5

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	39	41	33
Total number of relapses	17	18	38
Severity, n(%)*			
Mild	7 (41.2)	3 (16.7)	10 (26.3)
Moderate	7 (41.2)	14 (77.8)	25 (65.8)
Severe	3 (17.6)	1 (5.6)	3 (7.9)
Steroid use, n(%)*	12 (70.6)	17 (94.4)	31 (81.6)
Hospitalisations, n(%)*	4 (23.5)	5 (27.8)	13 (34.2)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
Date/time of run: 05SEP11 18:19 Page 1 of 1

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

Novartis

FTY720 HTA

HTA Table 15.4.4
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS >3.5

	FTY720D 1.25mg N=429	FTY720D 0.5mg N=425	Placebo N=418
Patients in group	3	7	4
Total number of relapses	6	4	8
Severity, n(%)*			
Mild	2 (33.3)	1 (25.0)	0 (0.0)
Moderate	2 (33.3)	3 (75.0)	7 (87.5)
Severe	2 (33.3)	0 (0.0)	1 (12.5)
Steroid use, n(%)*	6 (100.0)	4 (100.0)	6 (75.0)
Hospitalisations, n(%)*	2 (33.3)	1 (25.0)	6 (75.0)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.4.5
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Males

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	10	7	7
Total number of relapses	1	2	2
Severity, n(%)*			
Mild	0 (0.0)	2 (100.0)	0 (0.0)
Moderate	0 (0.0)	0 (0.0)	1 (50.0)
Severe	1 (100.0)	0 (0.0)	1 (50.0)
Steroid use, n(%)*	1 (100.0)	1 (50.0)	2 (100.0)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.4.6
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Females

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	18	20	23
Total number of relapses	4	4	7
Severity, n(%)*			
Mild	1 (25.0)	0 (0.0)	3 (42.9)
Moderate	3 (75.0)	4 (100.0)	4 (57.1)
Steroid use, n(%)*	3 (75.0)	2 (50.0)	4 (57.1)
Hospitalisations, n(%)*	1 (25.0)	0 (0.0)	1 (14.3)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.4.7

MS relapse (confirmed relapses only) characteristics by treatment

CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)

Baseline EDSS 0.0-3.5

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	26	26	25
Total number of relapses	5	5	6
Severity, n(%)*			
Mild	1 (20.0)	2 (40.0)	1 (16.7)
Moderate	3 (60.0)	3 (60.0)	4 (66.7)
Severe	1 (20.0)	0 (0.0)	1 (16.7)
Steroid use, n(%)*	4 (80.0)	2 (40.0)	4 (66.7)
Hospitalisations, n(%)*	1 (20.0)	0 (0.0)	0 (0.0)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 15.4.8
 MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS >3.5

	FTY720D 1.25mg N=420	FTY720D 0.5mg N=429	Interferon beta-1a i.m. N=431
Patients in group	2	1	5
Total number of relapses	0 (0.0)	1	3
Severity, n(%)*			
Mild	0 (0.0)	0 (0.0)	2 (66.7)
Moderate	0 (0.0)	1 (100.0)	1 (33.3)
Steroid use, n(%)*	0 (0.0)	1 (100.0)	2 (66.7)
Hospitalisations, n(%)*	0 (0.0)	0 (0.0)	1 (33.3)

* Percentage is calculated using total number of relapses as the denominator.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 15 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:19 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.1.1
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population
 Males

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.157	0.087	0.008
Steroid use*	0.257	0.234	1.000
Hospitalization*	0.039	0.048	1.000

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.1.2
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population
 Females

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.404	0.660	0.605
Steroid use*	0.872	0.883	1.000
Hospitalization*	0.144	0.736	0.344

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.1.3
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population
 Baseline EDSS 0.0-3.5

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.283	0.632	0.122
Steroid use*	0.558	0.694	0.874
Hospitalization*	0.550	0.915	0.687

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.1.4
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population
 Baseline EDSS >3.5

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.720	0.512	0.299
Steroid use*	0.337	0.687	1.000
Hospitalization*	0.578	0.196	0.095

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.1.5
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population
 Males

	FTY720D 1.25mg vs. Interferon beta-1a i.m. P-value	FTY720D 0.5mg vs. Interferon beta-1a i.m. P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.238	0.378	0.942
Steroid use*	1.000	0.154	0.134
Hospitalization*	0.244	0.546	1.000

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.1.6
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population
 Females

	FTY720D 1.25mg vs. Interferon beta-1a i.m. P-value	FTY720D 0.5mg vs. Interferon beta-1a i.m. P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.842	0.378	0.668
Steroid use*	0.850	0.249	0.401
Hospitalization*	0.854	0.220	0.234

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.1.7
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population
 Baseline EDSS 0.0-3.5

	FTY720D 1.25mg vs. Interferon beta-1a i.m. P-value	FTY720D 0.5mg vs. Interferon beta-1a i.m. P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.081	0.068	0.500
Steroid use*	0.862	0.111	0.120
Hospitalization*	1.000	0.104	0.156

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.1.8
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population
 Baseline EDSS >3.5

	FTY720D 1.25mg vs. Interferon beta-1a i.m. P-value	FTY720D 0.5mg vs. Interferon beta-1a i.m. P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.509	0.841	1.000
Steroid use*	0.671	0.425	1.000
Hospitalization*	0.204	0.551	0.698

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.2.1
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Males

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.829	0.585	1.000
Steroid use*	1.000	0.307	0.559
Hospitalization*	0.660	0.653	1.000

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.2.2
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Females

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.617	0.627	0.626
Steroid use*	1.000	0.177	0.669
Hospitalization*	1.000	0.529	0.669

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.2.3
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS 0.0-3.5

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.557	0.258	0.432
Steroid use*	0.210	1.000	0.380
Hospitalization*	0.195	0.521	0.433

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.2.4
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS >3.5

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.200	1.000	0.556
Steroid use*	0.036	1.000	1.000
Hospitalization*	0.079	1.000	1.000

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.2.5
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Males

	FTY720D 1.25mg vs. Interferon beta-1a i.m. P-value	FTY720D 0.5mg vs. Interferon beta-1a i.m. P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.910	0.578	0.306
Steroid use*	1.000	1.000	1.000
Hospitalization*	0.631	0.296	1.000

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.2.6
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Females

	FTY720D 1.25mg vs. Interferon beta-1a i.m. P-value	FTY720D 0.5mg vs. Interferon beta-1a i.m. P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.697	0.743	0.710
Steroid use*	1.000	0.498	0.723
Hospitalization*	0.593	1.000	0.523

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.2.7
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS 0.0-3.5

	FTY720D 1.25mg vs. Interferon beta-1a i.m. P-value	FTY720D 0.5mg vs. Interferon beta-1a i.m. P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.269	1.000	0.386
Steroid use*	0.761	0.503	0.266
Hospitalization*	0.428	0.722	0.185

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.2.8
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9)
 Baseline EDSS >3.5

	FTY720D 1.25mg vs. Interferon beta-1a i.m. P-value	FTY720D 0.5mg vs. Interferon beta-1a i.m. P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.414	0.276	0.640
Steroid use*	0.478	0.478	1.000
Hospitalization*	0.317	0.640	0.590

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.3.1
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Males

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.335	0.805	0.731
Steroid use*	1.000	0.411	1.000
Hospitalization*	1.000	1.000	1.000

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.3.2
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Females

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.380	1.000	0.402
Steroid use*	0.587	0.086	1.000
Hospitalization*	1.000	0.326	1.000

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.3.3
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS 0.0-3.5

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.801	0.480	1.000
Steroid use*	0.216	0.752	0.108
Hospitalization*	0.067	0.199	0.493

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.3.4
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS >3.5

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.029	1.000	0.167
Steroid use*	0.015	1.000	1.000
Hospitalization*	0.152	1.000	1.000

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.3.5
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Males

	FTY720D 1.25mg vs. Interferon beta-1a i.m. P-value	FTY720D 0.5mg vs. Interferon beta-1a i.m. P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.913	0.782	0.487
Steroid use*	0.605	0.630	0.540
Hospitalization*	0.605	0.510	1.000

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.3.6
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Females

	FTY720D 1.25mg vs. Interferon beta-1a i.m. P-value	FTY720D 0.5mg vs. Interferon beta-1a i.m. P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.658	0.830	1.000
Steroid use*	1.000	0.203	0.328
Hospitalization*	0.784	1.000	1.000

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.3.7
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS 0.0-3.5

	FTY720D 1.25mg vs. Interferon beta-1a i.m. P-value	FTY720D 0.5mg vs. Interferon beta-1a i.m. P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.514	0.656	0.822
Steroid use*	0.578	0.216	0.090
Hospitalization*	0.776	0.719	0.473

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.3.8
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (IFN non-responder: equal or more relapses in year-1 than year-2)
 Baseline EDSS >3.5

	FTY720D 1.25mg vs. Interferon beta-1a i.m. P-value	FTY720D 0.5mg vs. Interferon beta-1a i.m. P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.701	0.680	1.000
Steroid use*	0.487	0.478	1.000
Hospitalization*	0.288	0.640	0.580

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.4.1
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Males

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.017	0.280	0.279
Steroid use*	0.288	1.000	0.245
Hospitalization*	0.099	0.659	0.326

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.4.2
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Females

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.303	0.299	0.016
Steroid use*	0.702	0.236	0.569
Hospitalization*	1.000	0.496	0.662

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.4.3
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS 0.0-3.5

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.207	0.796	0.121
Steroid use*	0.482	0.414	0.088
Hospitalization*	0.536	0.763	1.000

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.4.4
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2301 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS >3.5

	FTY720D 1.25mg vs. Placebo P-value	FTY720D 0.5mg vs. Placebo P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	0.119	0.576	0.714
Steroid use*	0.473	0.515	0.714
Hospitalization*	0.277	0.222	1.000

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.4.5
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Males

	FTY720D 1.25mg vs. Interferon beta-1a i.m. P-value	FTY720D 0.5mg vs. Interferon beta-1a i.m. P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	1.000	0.333	0.333
Steroid use*	1.000	1.000	1.000
Hospitalization*	1.000	1.000	1.000

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.4.6
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Females

	FTY720D 1.25mg vs. Interferon beta-1a i.m. P-value	FTY720D 0.5mg vs. Interferon beta-1a i.m. P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	1.000	0.236	1.000
Steroid use*	1.000	1.000	1.000
Hospitalization*	1.000	1.000	1.000

Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Novartis

FTY720 HTA

HTA Table 16.4.7
 Comparison of MS relapse (confirmed relapses only) characteristics by treatment
 CFTY720D2302 ITT population (Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline)
 Baseline EDSS 0.0-3.5

	FTY720D 1.25mg vs. Interferon beta-1a i.m. P-value	FTY720D 0.5mg vs. Interferon beta-1a i.m. P-value	FTY720D 1.25mg vs. FTY720D 0.5mg P-value
Severity*	1.000	1.000	1.000
Steroid use*	1.000	0.567	0.524
Hospitalization*	0.455	0.567	1.000

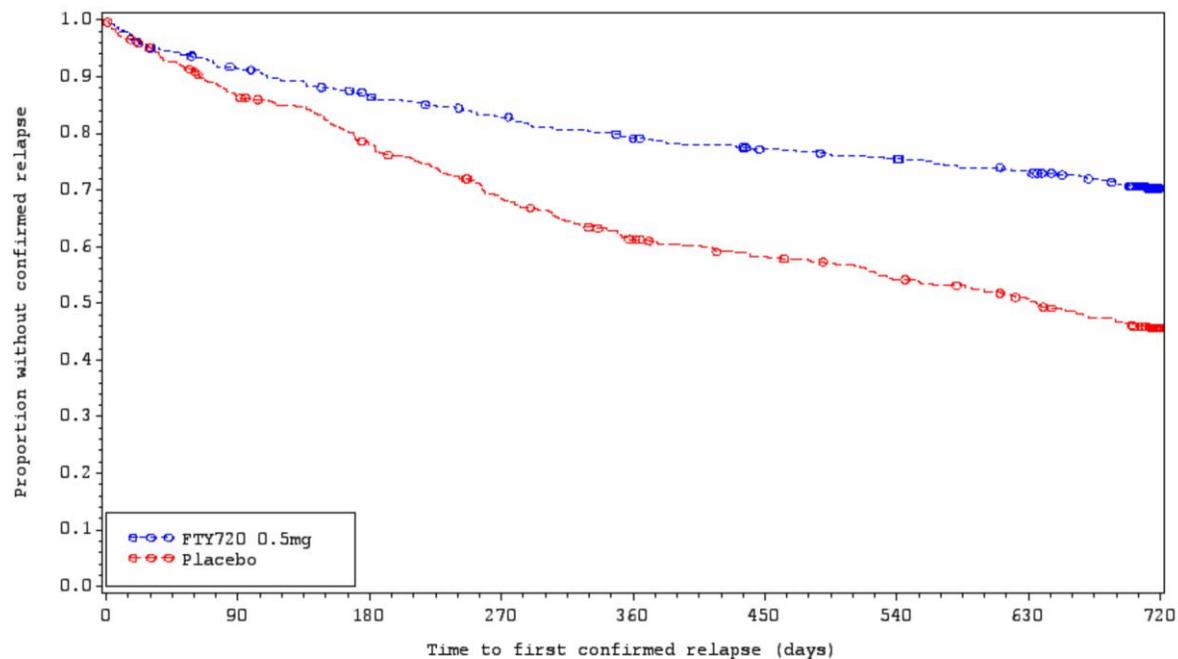
Note: comparison is at the relapse level, not patient level.
 * P-values are from Fisher's exact test.

SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA Tab 16 1.SAS status=FINAL
 Date/time of run: 05SEP11 18:22 Page 1 of 1

Anhang 4-I: Kaplan-Meier Darstellungen

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

HTA Figure 1.1: Kaplan-Meier plot of time to first confirmed relapse up to Month 24
 CFTY720D2301 ITT population

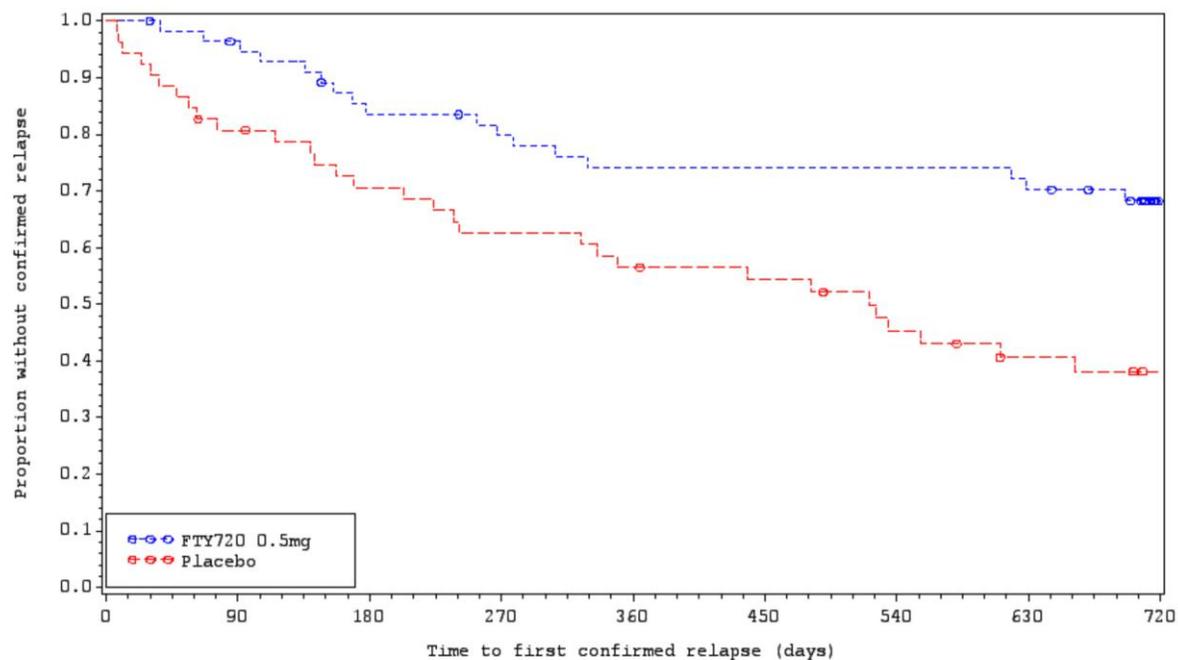


Time (days)	0	90	180	270	360	450	540	630	720
FTY720 0.5 mg	425	384	360	341	324	311	303	291	219
Placebo	418	356	316	276	242	226	208	190	145

Note: o = Censored observations
 SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_1_1.SAS status=FINAL
 Date/time of run: 16SEP11 22:00

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

HTA Figure 1.2: Kaplan-Meier plot of time to first confirmed relapse up to Month 24
 CF7Y720D2301 ITT population
 IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9

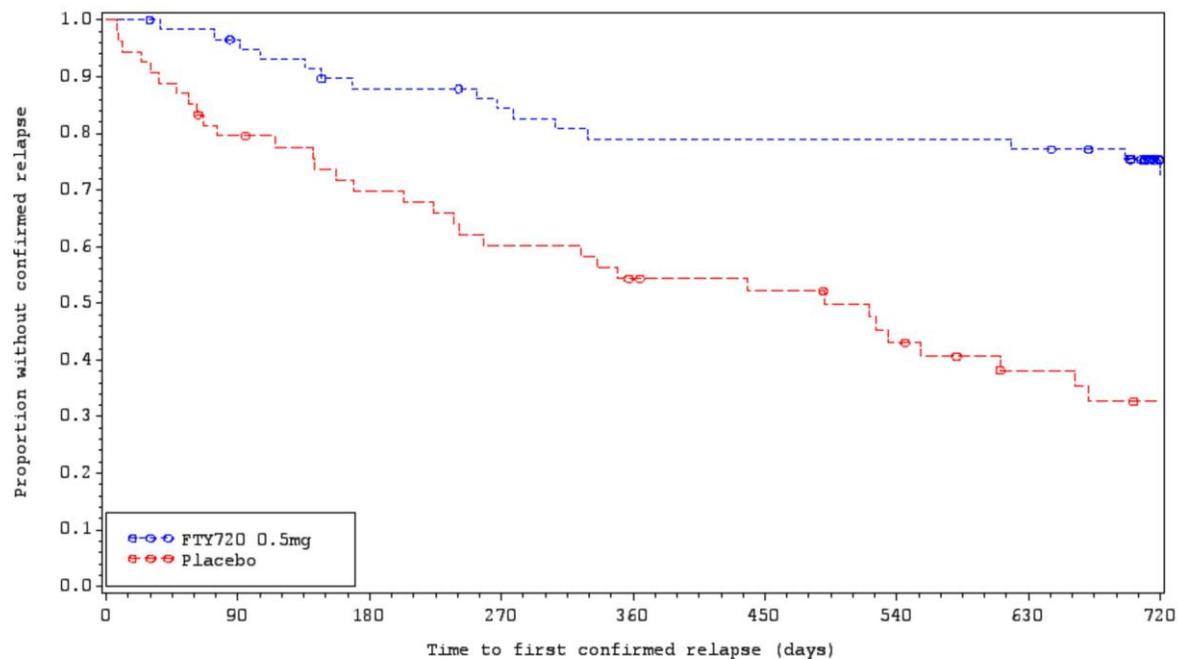


Time (days)	0	90	180	270	360	450	540	630	720
FTY720 0.5 mg	57	53	45	42	39	39	39	37	23
Placebo	52	41	35	31	28	25	20	16	13

Note: o = Censored observations
 SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_1_1.SAS status=FINAL
 Date/time of run: 16SEP11 22:00

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

HTA Figure 1.3: Kaplan-Meier plot of time to first confirmed relapse up to Month 24
 CFTY720D2301 ITT population
 IFN non-responder: equal or more relapses in year-1 than year-2

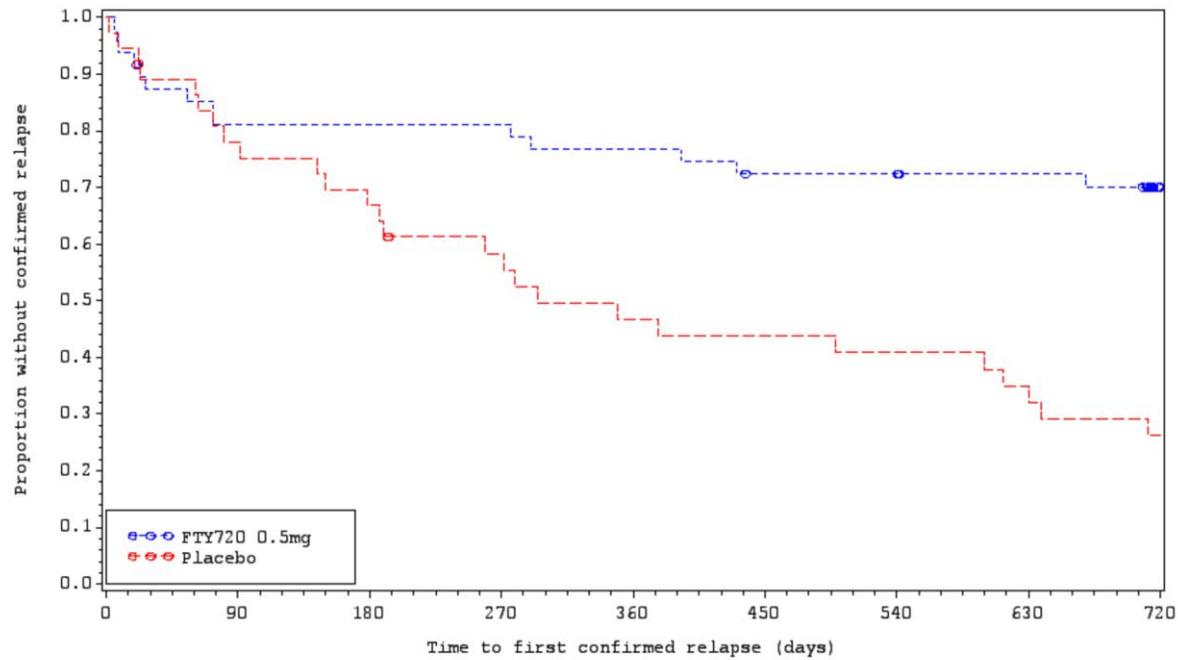


Pts at Risk:									
Time (days)	0	90	180	270	360	450	540	630	720
FTY720 0.5 mg	60	56	50	47	44	44	44	43	29
Placebo	54	42	36	31	27	24	19	14	11

Note: o = Censored observations
 SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_1_1.SAS status=FINAL
 Date/time of run: 16SEP11 22:00

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

HTA Figure 1.4: Kaplan-Meier plot of time to first confirmed relapse up to Month 24
 CFTY720D2301 ITT population
 Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline

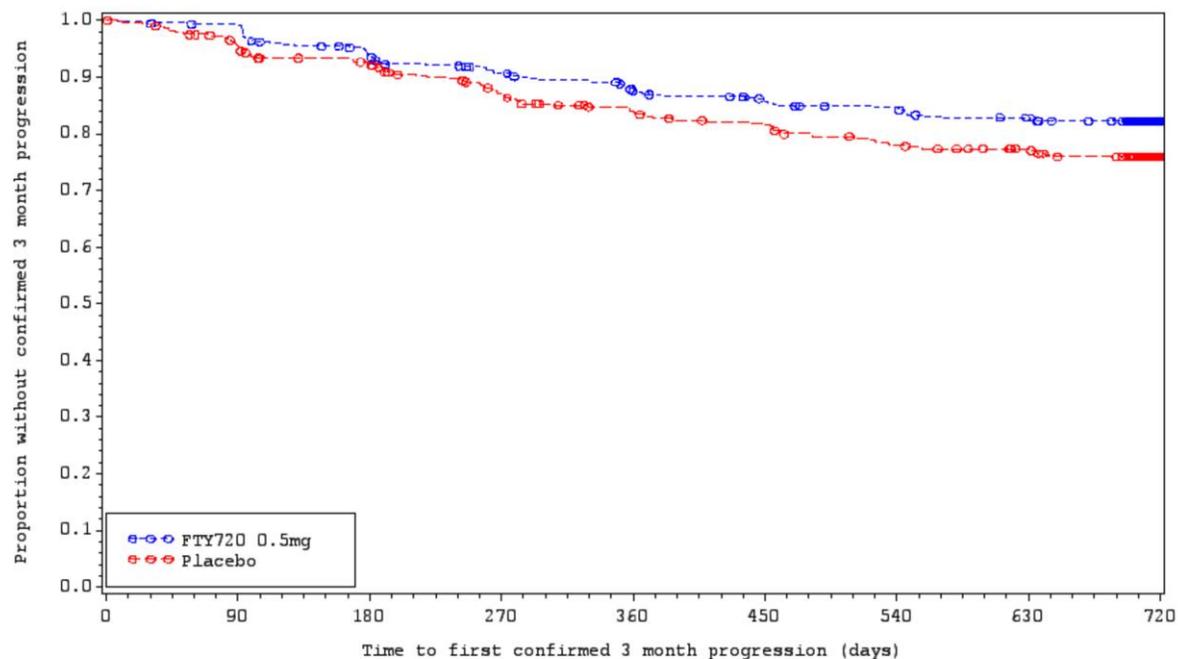


Pts at Risk:		0	90	180	270	360	450	540	630	720
Time (days)										
FTY720 0.5 mg		48	38	38	38	36	33	33	31	22
Placebo		37	28	24	20	16	15	14	12	9

Note: o = Censored observations
 SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_1_1.SAS status=FINAL
 Date/time of run: 16SEP11 22:00

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

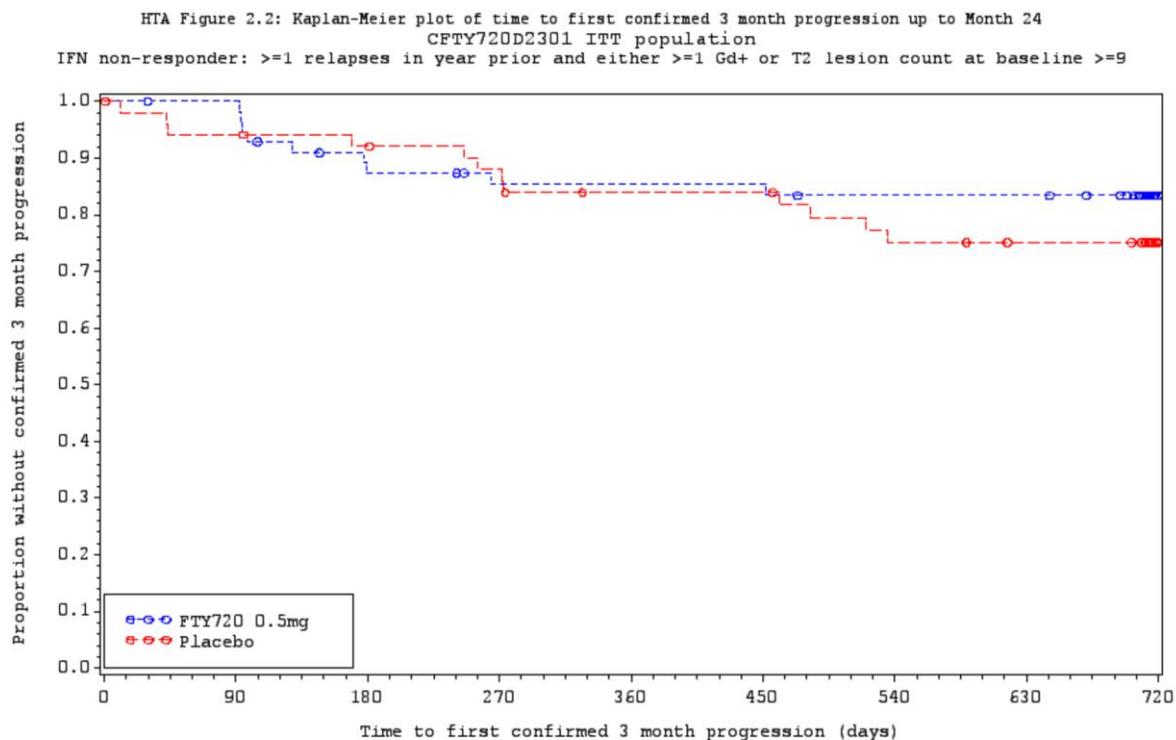
HTA Figure 2.1: Kaplan-Meier plot of time to first confirmed 3 month progression up to Month 24
 CFTY720D2301 ITT population



Pts at Risk:	0	90	180	270	360	450	540	630	720
Time (days)									
FTY720 0.5 mg	425	416	388	370	354	340	332	321	152
Placebo	418	391	371	341	320	308	290	279	143

Note: o = Censored observations
 SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_1_1.SAS status=FINAL
 Date/time of run: 16SEP11 22:00

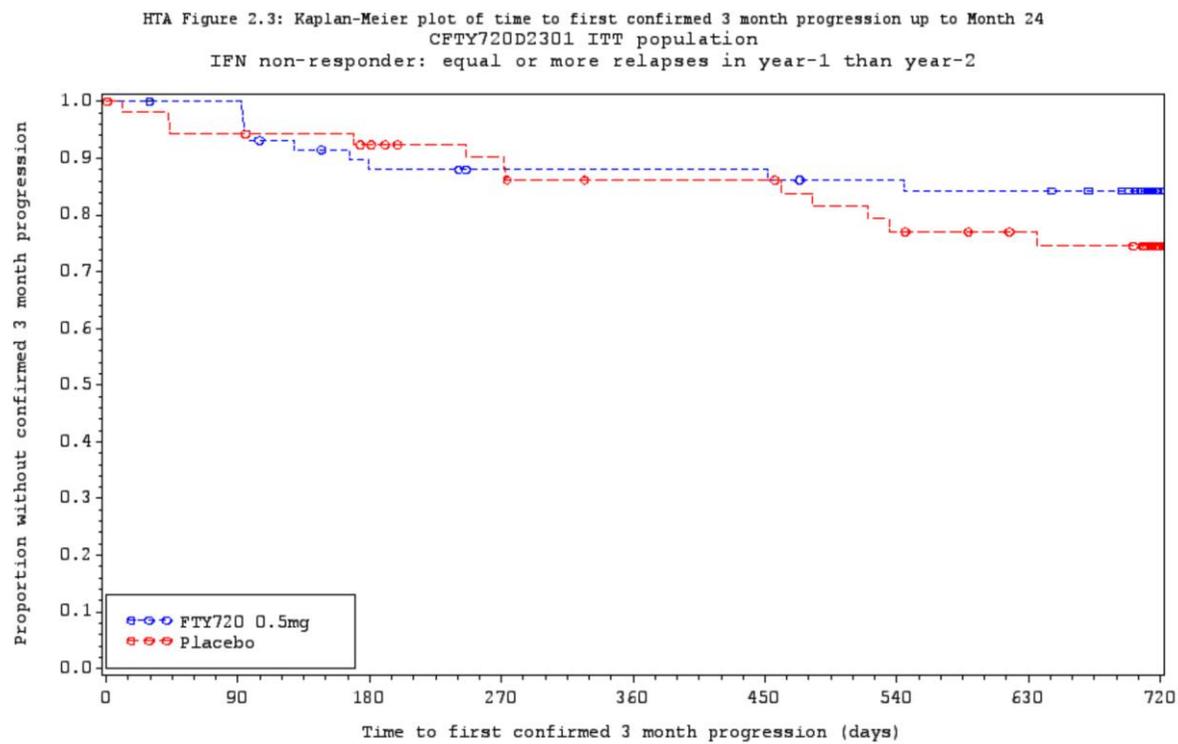
Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen



Pts at Risk:	0	90	180	270	360	450	540	630	720
Time (days)									
FTY720 0.5 mg	57	56	47	44	44	44	42	42	23
Placebo	52	48	46	43	39	39	34	32	18

Note: o = Censored observations
 SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_1_1.SAS status=FINAL
 Date/time of run: 16SEP11 22:00

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

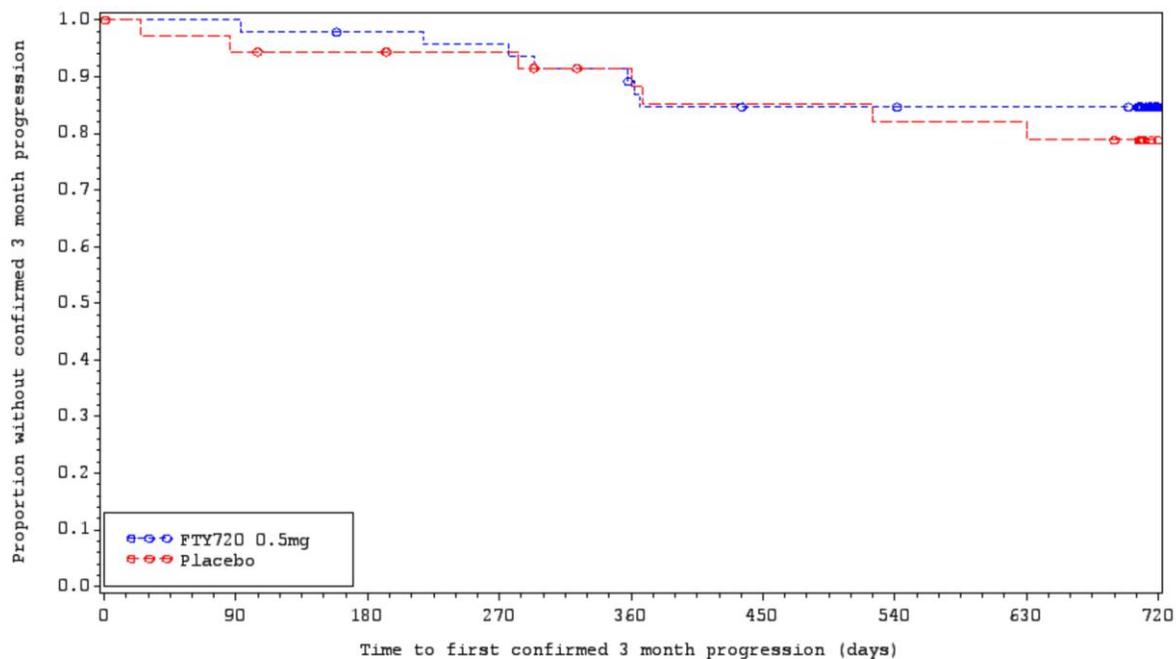


Pts at Risk:		0	90	180	270	360	450	540	630	720
Time (days)		0	90	180	270	360	450	540	630	720
FTY720 0.5 mg		60	59	50	48	48	48	46	45	23
Placebo		54	50	47	43	39	39	34	31	18

Note: ○ = Censored observations
 SOURCE: C:\Humerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_1_1.SAS status=FINAL
 Date/Time of run: 12SEP11 22:00

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

HTA Figure 2.4: Kaplan-Meier plot of time to first confirmed 3 month progression up to Month 24
 CFTY720D2301 ITT population
 Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline

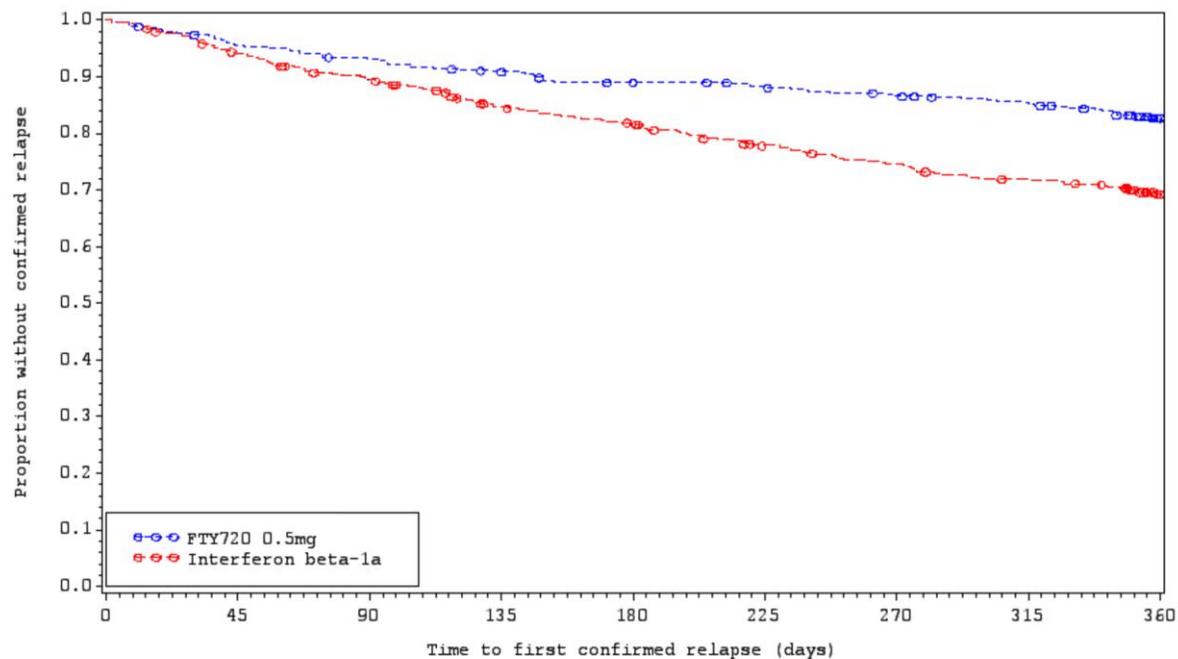


Pts at Risk:		0	90	180	270	360	450	540	630	720
Time (days)										
FTY720 0.5 mg		48	47	45	44	40	37	37	36	17
Placebo		37	34	33	32	29	27	26	26	16

Note: o = Censored observations
 SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_1_1.SAS status=FINAL
 Date/time of run: 16SEP11 22:00

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

HTA Figure 4.1: Kaplan-Meier plot of time to first confirmed relapse up to Month 12
 CFTY720D2302 ITT population

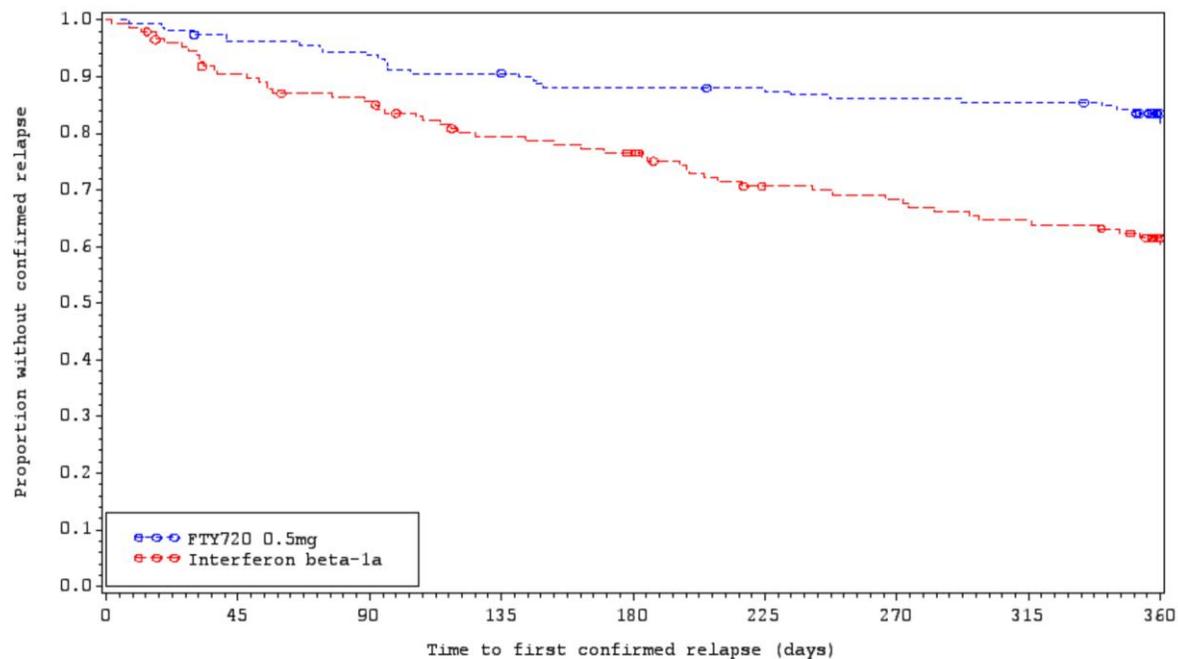


Time (days)	0	45	90	135	180	225	270	315	360
FTY720 0.5 mg	429	408	397	385	374	368	359	351	294
IFN beta-1a im	431	402	379	350	335	313	299	286	240

Note: o = Censored observations
 SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_4_1.SAS status=FINAL
 Date/time of run: 16SEP11 22:03

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

HTA Figure 4.2: Kaplan-Meier plot of time to first confirmed relapse up to Month 12
 CFTY720D2302 ITT population
 IFN non-responder: >=1 relapses in year prior and either >=1 Gd+ or T2 lesion count at baseline >=9

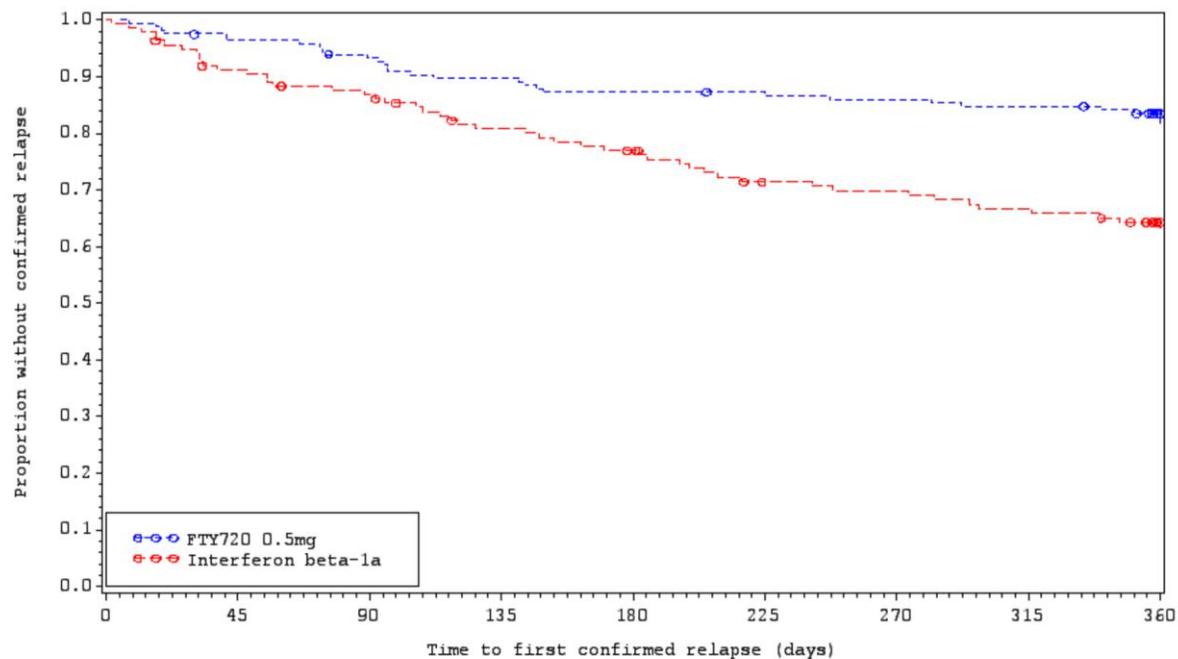


Time (days)	0	45	90	135	180	225	270	315	360
FTY720 0.5 mg	160	153	149	144	139	138	135	134	113
IFN beta-1a im	149	132	124	112	107	94	91	86	72

Note: o = Censored observations
 SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_4_1.SAS status=FINAL
 Date/time of run: 16SEP11 22:03

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

HTA Figure 4.3: Kaplan-Meier plot of time to first confirmed relapse up to Month 12
 CFTY720D2302 ITT population
 IFN non-responder: equal or more relapses in year-1 than year-2

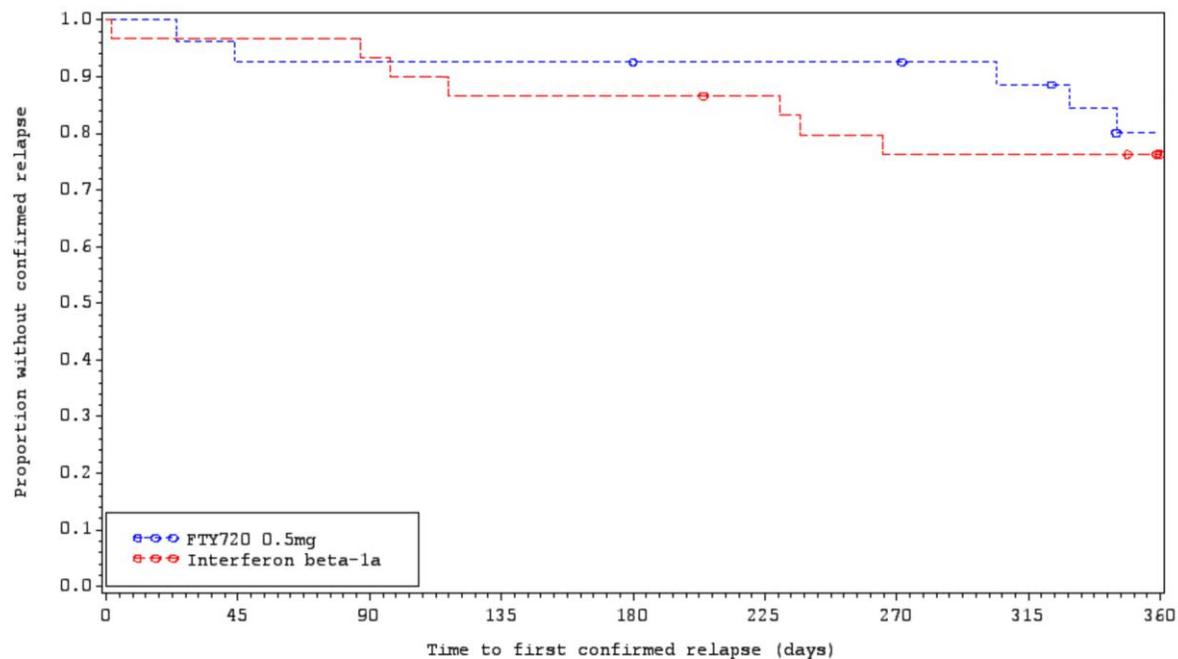


Time (days)	0	45	90	135	180	225	270	315	360
FTY720 0.5 mg	166	159	153	147	143	142	140	138	118
IFN beta-1a im	138	124	117	106	100	89	87	83	71

Note: o = Censored observations
 SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_4_1.SAS status=FINAL
 Date/time of run: 16SEP11 22:03

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

HTA Figure 4.4: Kaplan-Meier plot of time to first confirmed relapse up to Month 12
 CFTY720D2302 ITT population
 Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline

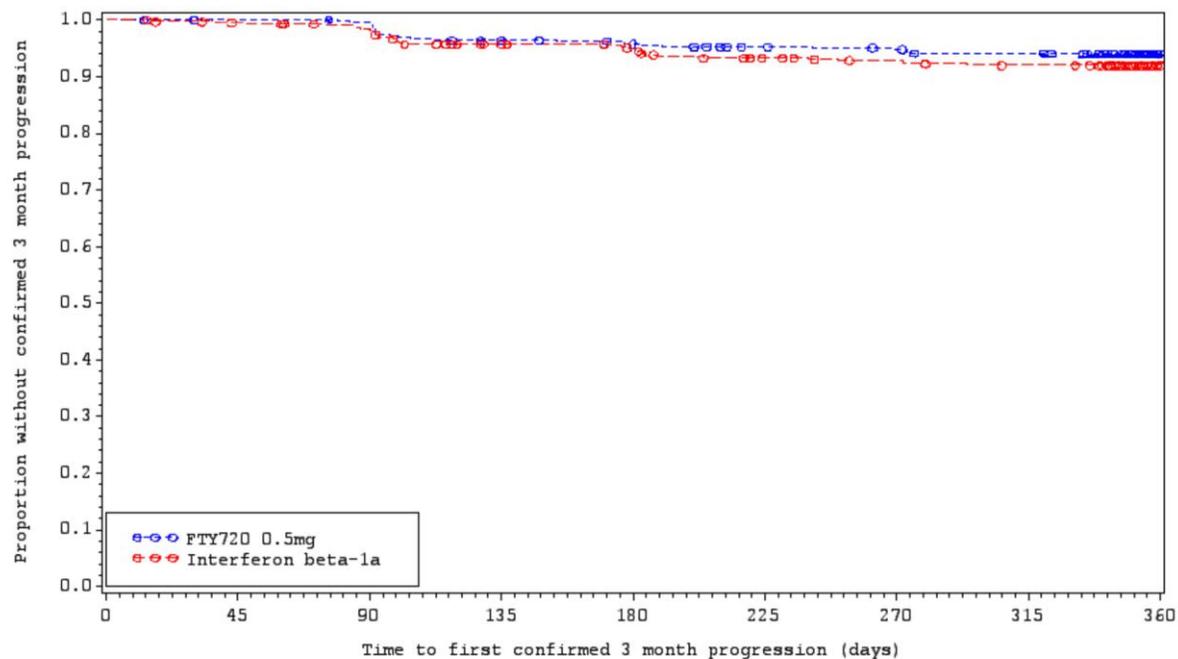


Pts at Risk:		0	45	90	135	180	225	270	315	360
Time (days)										
FTY720 0.5 mg	27	25	25	25	25	25	24	24	22	18
IFN beta-1a im	30	29	28	26	26	26	25	22	22	19

Note: o = Censored observations
 SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_4_1.SAS status=FINAL
 Date/time of run: 16SEP11 22:03

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

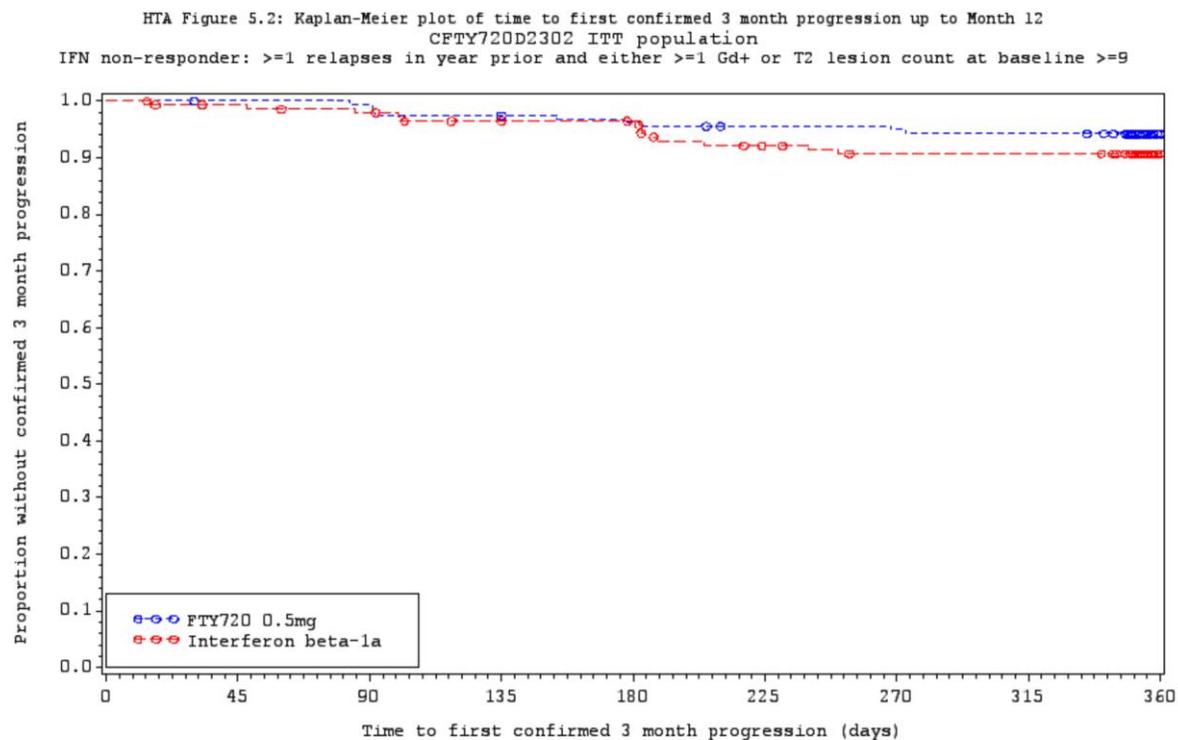
HTA Figure 5.1: Kaplan-Meier plot of time to first confirmed 3 month progression up to Month 12
 CFTY720D2302 ITT population



Pts at Risk:		0	45	90	135	180	225	270	315	360
Time (days)										
FTY720 0.5 mg		429	427	424	409	403	395	391	386	225
IFN beta-1a im		431	425	417	397	390	374	368	363	202

Note: o = Censored observations
 SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_4_1.SAS status=FINAL
 Date/time of run: 16SEP11 22:03

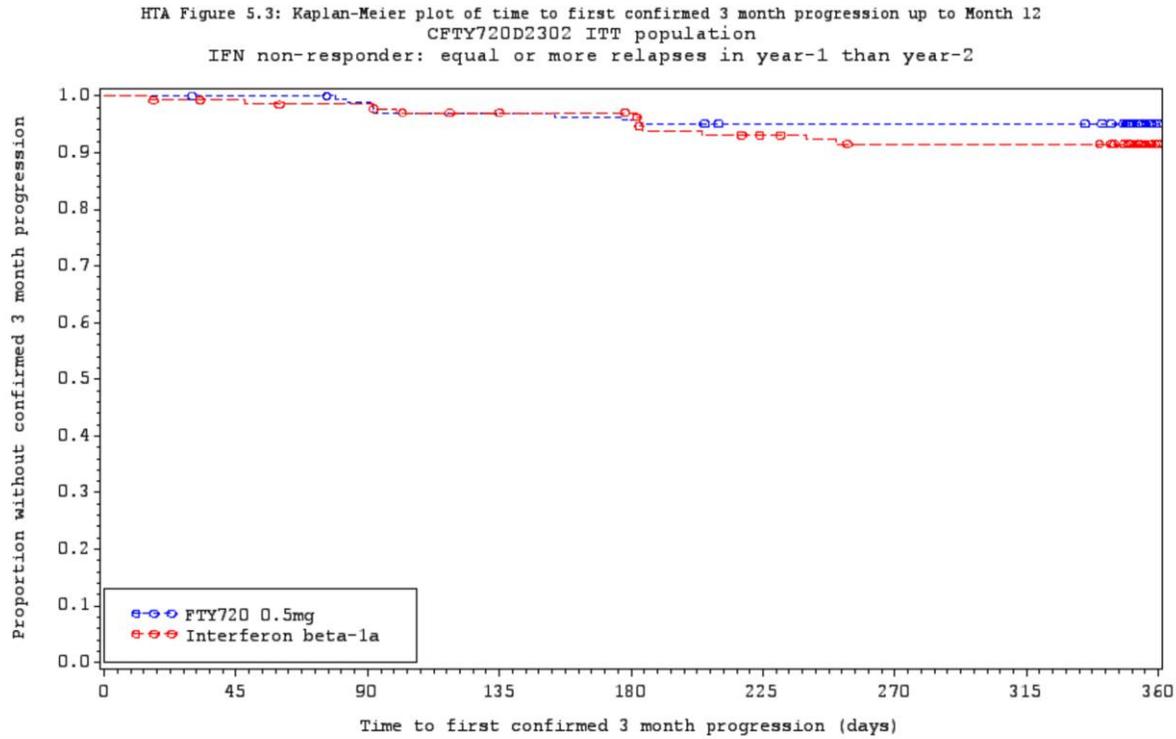
Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen



Pts at Risk:	0	45	90	135	180	225	270	315	360
Time (days)									
FTY720 0.5 mg	160	159	158	155	152	149	148	147	90
IFN beta-1a im	149	145	142	137	135	123	119	119	63

Note: o = Censored observations
 SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_4_1.SAS status=FINAL
 Date/time of run: 16SEP11 22:03

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

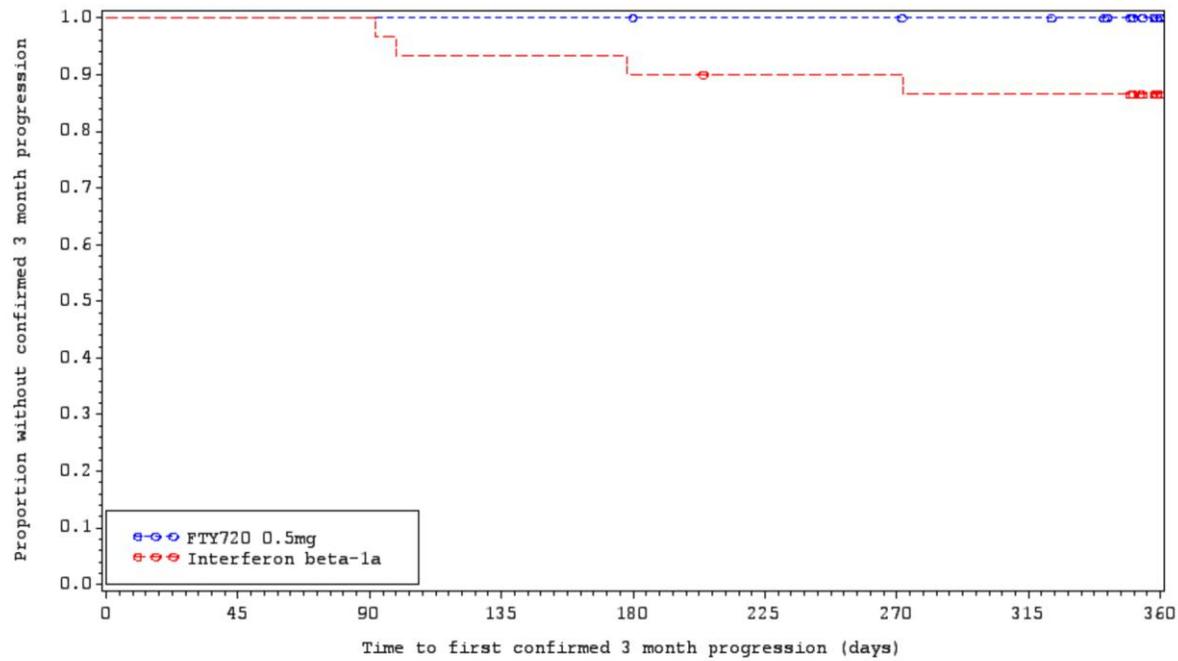


Pts at Risk:		0	45	90	135	180	225	270	315	360
Time (days)										
FTY720 0.5 mg	166	165	162	159	157	154	154	154	154	96
IFN beta-1a im	138	135	133	128	126	116	116	112	112	62

Note: o = Censored observations
 SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_4_1.SAS status=FINAL
 Date/time of run: 16SEP11 22:03

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

HTA Figure 5.4: Kaplan-Meier plot of time to first confirmed 3 month progression up to Month 12
 CFTY720D2302 ITT population
 Rapidly evolving: IFN naive, >=2 relapses in year-1 and >=1 Gd+ at baseline



Pts at Risk:		0	45	90	135	180	225	270	315	360
Time (days)										
FTY720 0.5 mg	27	27	27	27	27	27	26	26	25	16
IFN beta-1a im	30	30	30	28	27	26	26	26	25	16

Note: o = Censored observations
 SOURCE: C:\Numerus\Studies\Nov\Nov101\Analysis\Prod\Progs\HTA_Fig_4_1.SAS status=FINAL
 Date/time of run: 16SEP11 22:03