

Dossier zur Nutzenbewertung gemäß § 35a SGB V

*THC/CBD (Sativex[®] Spray zur Anwendung in der
Mundhöhle)*

Almirall Hermal GmbH

Modul 4A

Zur Symptomverbesserung bei Patienten mit mittelschwerer bis schwerer Spastik aufgrund von Multipler Sklerose (MS), die nicht angemessen auf eine andere anti-spastische Arzneimitteltherapie angesprochen haben und die eine klinisch erhebliche Verbesserung von mit der Spastik verbundenen Symptomen während eines Anfangstherapieversuchs aufzeigen.

**Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen**

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	4
Abbildungsverzeichnis	9
Abkürzungsverzeichnis	10
4 Modul 4 – allgemeine Informationen	12
4.1 Zusammenfassung der Inhalte von Modul 4.....	13
4.2 Methodik	16
4.2.1 Fragestellung	16
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung	17
4.2.3 Informationsbeschaffung	19
4.2.3.1 Studien des pharmazeutischen Unternehmers	20
4.2.3.2 Bibliografische Literaturrecherche	20
4.2.3.3 Suche in Studienregistern	21
4.2.3.4 Selektion relevanter Studien	22
4.2.4 Bewertung der Aussagekraft der Nachweise.....	22
4.2.5 Informationssynthese und -analyse	23
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien	23
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	24
4.2.5.3 Meta-Analysen.....	27
4.2.5.4 Sensitivitätsanalysen	28
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	29
4.2.5.6 Indirekte Vergleiche	30
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen	31
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	32
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel	32
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	32
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	36
4.3.1.1.3 Studien aus der Suche in Studienregistern	39
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	39
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	41
4.3.1.2.1 Studiendesign und Studienpopulationen	41
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	51
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	52
4.3.1.3.1 <Endpunkt xxx> – RCT	53
4.3.1.3.1.1 Absolute Veränderung des Grads der Spastizität auf der NRS – RCT	53
4.3.1.3.1.2 Anzahl Patienten mit klinisch relevantem Rückgang der Schwere der Spastizität (Responder) auf der NRS – RCT	59
4.3.1.3.1.3 Anzahl Spastiken pro Tag – RCT.....	67

4.3.1.3.1.4	Schlafqualität – RCT	71
4.3.1.3.1.5	Gesundheitszustand (EQ-5D) – RCT	75
4.3.1.3.1.6	Lebensqualität – RCT	79
4.3.1.3.1.7	Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC)– RCT	83
4.3.1.3.1.8	Globalbeurteilung der klinischen Veränderung durch die betreuende Person (CGIC) – RCT	87
4.3.1.3.1.9	Globalbeurteilung der klinischen Veränderung durch den Arzt (PhGIC) – RCT	91
4.3.1.3.1.10	Aktivitäten des täglichen Lebens (Barthel Index) – RCT	94
4.3.1.3.1.11	Gehfähigkeit (10 m-Gehstreckentest) – RCT	98
4.3.1.3.1.12	Nebenwirkungen und schwerwiegende Nebenwirkungen (mit möglichem Zusammenhang zur Studienmedikation) – RCT	102
4.3.1.3.2	Subgruppenanalysen – RCT	105
4.3.1.3.3	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien	106
4.3.2	Weitere Unterlagen	111
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	111
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	111
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche	111
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	112
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT	112
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	114
4.3.2.2	Nicht randomisierte vergleichende Studien	115
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	115
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien	115
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	116
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien	116
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	117
4.3.2.3	Weitere Untersuchungen	117
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	117
4.3.2.3.2	Studien aus der bibliografischen Literaturrecherche	123
4.3.2.3.3	Studien aus der Suche in Studienregistern	124
4.3.2.3.4	Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel	124
4.3.2.3.5	Charakteristika der weiteren Untersuchungen	125
4.3.2.3.6	Ergebnisse aus weiteren Untersuchungen	125
4.3.2.3.6.1	<Endpunkt xxx> – weitere Untersuchungen	125
4.3.2.3.6.2	Subgruppenanalysen – weitere Untersuchungen	135
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen	135
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens	136
4.4.1	Beurteilung der Aussagekraft der Nachweise	136
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß	139

4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	142
4.4.4	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens....	142
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	143
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	143
4.5.2	Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	144
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	144
4.5.4	Verwendung von Surrogatendpunkten	144
4.6	Liste der eingeschlossenen Studien.....	145
4.7	Referenzliste.....	146
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		150
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....		154
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund		157
Anhang 4-D : Liste der abgebrochenen Studien		158
Anhang 4-E : Liste der laufenden Studien.....		159
Anhang 4-F : Methodik der eingeschlossenen Studien – RCT.....		160
Anhang 4-G : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten		199

Tabellenverzeichnis

	Seite
Tabelle 4-1: Patientenrelevante Endpunkte.....	18
Tabelle 4-2: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	33
Tabelle 4-3: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	36
Tabelle 4-4: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	39
Tabelle 4-5: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	40
Tabelle 4-6: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	42
Tabelle 4-7: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	45
Tabelle 4-8: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	47
Tabelle 4-9: Medikamentöse Vorbehandlung und Cannabis-Konsum	50
Tabelle 4-10: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	51
Tabelle 4-11: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	52
Tabelle 4-12: Operationalisierung von Absolute Veränderung des Grads der Spastizität auf der NRS.....	54
Tabelle 4-13: Bewertung des Verzerrungspotenzials für Absolute Veränderung des Grads der Spastizität auf der NRS in RCT mit dem zu bewertenden Arzneimittel.....	55
Tabelle 4-14: Ergebnisse für Absolute Veränderung des Grads der Spastizität auf der NRS aus RCT mit dem zu bewertenden Arzneimittel	56
Tabelle 4-15: PP-Analyse der Studie GWCL0403 für Absolute Veränderung des Grads der Spastizität auf der NRS	57
Tabelle 4-16: Operationalisierung von Anzahl Patienten mit klinisch relevantem Rückgang der Schwere der Spastizität (Responder)	60
Tabelle 4-17: Bewertung des Verzerrungspotenzials für Anzahl Patienten mit klinisch relevantem Rückgang der Schwere der Spastizität (Responder) in RCT mit dem zu bewertenden Arzneimittel	61
Tabelle 4-18: Ergebnisse für Anzahl Patienten mit klinisch relevantem Rückgang der Schwere der Spastizität (Responder, 30% Verminderung) gemessen mittels NRS aus RCT mit dem zu bewertenden Arzneimittel	62
Tabelle 4-19: PP-Analyse der Studie GWCL0403 für Anzahl Patienten mit klinisch relevantem Rückgang der Schwere der Spastizität gemessen mittels NRS	63
Tabelle 4-20: Operationalisierung von Anzahl Spastiken pro Tag	67

Tabelle 4-21: Bewertung des Verzerrungspotenzials für Anzahl Spastiken pro Tag in RCT mit dem zu bewertenden Arzneimittel	68
Tabelle 4-22: Ergebnisse für Anzahl Spastiken pro Tag aus RCT mit dem zu bewertenden Arzneimittel.....	69
Tabelle 4-23: Operationalisierung von Schlafqualität	71
Tabelle 4-24: Bewertung des Verzerrungspotenzials für Schlafqualität in RCT mit dem zu bewertenden Arzneimittel	72
Tabelle 4-25: Ergebnisse für Schlafqualität aus RCT mit dem zu bewertenden Arzneimittel.....	73
Tabelle 4-26: Operationalisierung von Gesundheitszustand (EQ-5D)	75
Tabelle 4-27: Bewertung des Verzerrungspotenzials für den Gesundheitszustand(EQ-5D) in RCT mit dem zu bewertenden Arzneimittel	76
Tabelle 4-28: Ergebnisse für Gesundheitszustand (EQ-5D) aus RCT mit dem zu bewertenden Arzneimittel	77
Tabelle 4-29: Operationalisierung von Lebensqualität	79
Tabelle 4-30: Bewertung des Verzerrungspotenzials für Lebensqualität in RCT mit dem zu bewertenden Arzneimittel	80
Tabelle 4-31: Ergebnisse für Lebensqualität aus RCT mit dem zu bewertenden Arzneimittel*	81
Tabelle 4-32: Operationalisierung von Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC)	84
Tabelle 4-33: Bewertung des Verzerrungspotenzials für Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC) in RCT mit dem zu bewertenden Arzneimittel.....	84
Tabelle 4-34: Ergebnisse für Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC) aus RCT mit dem zu bewertenden Arzneimittel	85
Tabelle 4-35: Operationalisierung von Globalbeurteilung der klinischen Veränderung durch die betreuende Person (CGIC)	88
Tabelle 4-36: Bewertung des Verzerrungspotenzials für Globalbeurteilung der klinischen Veränderung durch die betreuende Person (CGIC) in RCT mit dem zu bewertenden Arzneimittel.....	88
Tabelle 4-37: Ergebnisse für Globalbeurteilung der klinischen Veränderung durch die betreuende Person (CGIC) aus RCT mit dem zu bewertenden Arzneimittel	89
Tabelle 4-38: Operationalisierung von Globalbeurteilung der klinischen Veränderung durch den Arzt (PhGIC)	91
Tabelle 4-39: Bewertung des Verzerrungspotenzials für Globalbeurteilung der klinischen Veränderung durch den Arzt (PhGIC) in RCT mit dem zu bewertenden Arzneimittel.....	92
Tabelle 4-40: Ergebnisse für Globalbeurteilung der klinischen Veränderung durch den Arzt (PhGIC) aus RCT mit dem zu bewertenden Arzneimittel	93
Tabelle 4-41: Operationalisierung von Aktivitäten des täglichen Lebens (Barthel Index).....	94
Tabelle 4-42: Bewertung des Verzerrungspotenzials für Aktivitäten des täglichen Lebens (Barthel Index)in RCT mit dem zu bewertenden Arzneimittel.....	95

Tabelle 4-43: Ergebnisse für Aktivitäten des täglichen Lebens (Barthel Index) aus RCT mit dem zu bewertenden Arzneimittel	96
Tabelle 4-44: Operationalisierung von Gehfähigkeit (10 m-Gehstreckentest)	98
Tabelle 4-45: Bewertung des Verzerrungspotenzials für Gehfähigkeit (10 m-Gehstreckentest)in RCT mit dem zu bewertenden Arzneimittel	99
Tabelle 4-46: Ergebnisse für Gehfähigkeit (10 m-Gehstreckentest)aus RCT mit dem zu bewertenden Arzneimittel	100
Tabelle 4-47: Operationalisierung von Nebenwirkungen	102
Tabelle 4-48: Bewertung des Verzerrungspotenzials für Nebenwirkungen mit möglichem Zusammenhang zur Studienmedikation in RCT mit dem zu bewertenden Arzneimittel	103
Tabelle 4-49: Ergebnisse für Nebenwirkungen aus RCT mit dem zu bewertenden Arzneimittel.....	104
Tabelle 4-50: Ergebnisse für Schwerwiegende Nebenwirkungen mit möglichem Zusammenhang zur Studienmedikation aus RCT mit dem zu bewertenden Arzneimittel	104
Tabelle 4-51: Zusammenfassung der Ergebnisse zur Verbesserung des Grades der Spastik auf der NRS.....	107
Tabelle 4-52: Zusammenfassung der Ergebnisse zur klinisch relevanten Verbesserung der Spastik auf der NRS	107
Tabelle 4-53: Zusammenfassung der Ergebnisse zur Anzahl der Spastiken	108
Tabelle 4-54: Zusammenfassende Einschätzung aller Endpunkte.....	110
Tabelle 4-55: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	112
Tabelle 4-56: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	112
Tabelle 4-57: Operationalisierung von <Endpunkt xxx>.....	113
Tabelle 4-58: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	113
Tabelle 4-59: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	114
Tabelle 4-60: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	116
Tabelle 4-61: Operationalisierung von <Endpunkt xxx>.....	116
Tabelle 4-62: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	117
Tabelle 4-63: Liste der Studien des pharmazeutischen Unternehmers – weitere Untersuchungen mit dem zu bewertenden Arzneimittel	119
Tabelle 4-64: Studien des pharmazeutischen Unternehmers, die nicht für weitere Untersuchungen mit dem zu bewertenden Arzneimittel in Frage kommen.....	122
Tabelle 4-65: Relevante Studien aus der Suche in Studienregistern – weitere Untersuchungen mit dem zu bewertenden Arzneimittel	124

Tabelle 4-66: Studienpool – Weitere Untersuchungen mit dem zu bewertenden Arzneimittel.....	124
Tabelle 4-67: Operationalisierung von Absolute Veränderung des Grads der Spastizität auf der NRS.....	125
Tabelle 4-68: Bewertung des Verzerrungspotenzials für Absolute Veränderung des Grads der Spastizität auf der NRS in RCT mit dem zu bewertenden Arzneimittel.....	126
Tabelle 4-69: Ergebnisse für Absolute Veränderung des Grads der Spastizität auf der NRS aus RCT mit dem zu bewertenden Arzneimittel	126
Tabelle 4-70: Operationalisierung von Schlafqualität	127
Tabelle 4-71: Bewertung des Verzerrungspotenzials für Schlafqualität in RCT mit dem zu bewertenden Arzneimittel	127
Tabelle 4-72: Ergebnisse für Schlafqualität aus RCT mit dem zu bewertenden Arzneimittel.....	128
Tabelle 4-73: Operationalisierung von Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC)	128
Tabelle 4-74: Bewertung des Verzerrungspotenzials für Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC) in RCT mit dem zu bewertenden Arzneimittel... ..	129
Tabelle 4-75: Ergebnisse für Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC) aus RCT mit dem zu bewertenden Arzneimittel	129
Tabelle 4-76: Operationalisierung von klinische Veränderung durch die betreuende Person (CGIC).....	130
Tabelle 4-77: Bewertung des Verzerrungspotenzials für Globalbeurteilung der klinischen Veränderung durch die betreuende Person (CGIC) in RCT mit dem zu bewertenden Arzneimittel.....	130
Tabelle 4-78: Ergebnisse für Globalbeurteilung der klinischen Veränderung durch die betreuende Person (CGIC) aus RCT mit dem zu bewertenden Arzneimittel	131
Tabelle 4-79: Operationalisierung von Gehfähigkeit.....	132
Tabelle 4-80: Bewertung des Verzerrungspotenzials für Gehfähigkeit (10 m-Gehstreckentest)in RCT mit dem zu bewertenden Arzneimittel	132
Tabelle 4-81: Ergebnisse für Gehfähigkeit (10 m-Gehstreckentest)aus RCT mit dem zu bewertenden Arzneimittel	133
Tabelle 4-82: Operationalisierung von Nebenwirkungen	133
Tabelle 4-83: Bewertung des Verzerrungspotenzials für Nebenwirkungen in RCT mit dem zu bewertenden Arzneimittel.....	134
Tabelle 4-84: Ergebnisse für Nebenwirkungen aus RCT mit dem zu bewertenden Arzneimittel.....	135
Tabelle 4-85: Ergebnisse für schwerwiegende Nebenwirkungen aus RCT mit dem zu bewertenden Arzneimittel	135
Tabelle 4-86: Zusammenfassende Einschätzung aller Endpunkte für die Studie GWSP0702.....	136

Tabelle 4-87: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	142
Tabelle 4-88: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens	143
Tabelle 4-89 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel	158
Tabelle 4-90 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel.....	159
Tabelle 4-91 (Anhang): Studiendesign und -methodik für Studie GWMS0001.....	161
Tabelle 4-92 (Anhang): Studiendesign und -methodik für Studie GWMS0106.....	170
Tabelle 4-93 (Anhang): Studiendesign und -methodik für Studie GWCL0403	178
Tabelle 4-94 (Anhang): Studiendesign und -methodik für Studie GWSP0604.....	186
Tabelle 4-95 (Anhang): Studiendesign und -methodik für Studie ARA2009.....	194
Tabelle 4-96 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GWMS0001	200
Tabelle 4-97 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GWMS0106	213
Tabelle 4-98 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GWCL0403.....	232
Tabelle 4-99 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GWSP0604.....	263
Tabelle 4-100 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie ARA2009.....	303

Abbildungsverzeichnis

	Seite
Abbildung 1: Einschlusskriterien für die Berücksichtigung von Studien in der Bewertung des Zusatznutzens.....	19
Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel	38
Abbildung 3: Art und Dauer der Behandlungsphasen und der Erhebungszeitpunkte der konfirmatorischen Wirksamkeitsstudie GWSP0604.....	49
Abbildung 4: Meta-Analyse für den Endpunkt Veränderung des Grads der Spastizität	58
Abbildung 5: Überprüfung der Homogenität für den Endpunkt „Klinisch relevanter Rückgang der Schwere der Spastizität“ (Responder, mind. 30 %).....	64
Abbildung 6: Meta-Analyse für den Endpunkt „Klinisch relevanter Rückgang der Schwere der Spastizität“ (Responder, mind. 30%) ohne die Studie GWMS0001	65
Abbildung 7: Meta-Analyse für den Endpunkt „Klinisch relevanter Rückgang der Schwere der Spastizität“ (Responder, mind. 30%) inklusive aller Patienten mit Spastik aus GWMS0001	65
Abbildung 8: Meta-Analyse für den Endpunkt „Klinisch relevanter Rückgang der Schwere der Spastizität“ (Responder, mind. 50 %) inklusive aller Patienten mit Spastik aus GWMS0001	66
Abbildung 9: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach weiteren Untersuchungen mit dem zu bewertenden Arzneimittel	123

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ANCOVA	Analysis of covariance
BI	Barthel Index
CBD	Cannabidiol
CGIC	Caregiver Global Impression of Change
95% CI	95% Confidence Interval
CONSORT	Consolidated Standards of Reporting Trials
DIMDI	Deutsches Institut für Medizinische Dokumentation
EG	Europäische Gemeinschaft
EMA	European Medicines Agency
EQ-5D	EQ-5D Fragebogen 3 level version
EudraCT	European Union Drug Regulating Authorities Clinical Trials
EuroQol	EuroQol Group
GCP	Good Clinical Practice
GIC	Global Impression of Change
ITT	Intention to treat
IVRS	Interactive Voice Response System
MS	Multiple Sklerose
MSQoL-54	Multiple Sclerosis Quality of Life-54 Instrument
MTC	Mixed Treatment Comparison
MW	Mittelwert
NA	Nicht zutreffend
ND	Nicht durchgeführt
NR	Nicht dokumentiert
NRS	Numeric rating scale
PhGIC	Physician Global Impression of Change
PP	Per Protocol
RCT	Randomized Controlled Trial
SD	Standard Deviation
SF-36	Short Form 36 Health Survey Questionnaire
SGB	Sozialgesetzbuch

SGIC	SubjectGlobal Impression of Change
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
THC	Tetrahydrocannabinol
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
VAS	Visual Analogue Scale
VRS	Verbal Rating Scale
WHO	World Health Organization

4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

- Zusammenfassung (Abschnitt 4.1)
- Angaben zur Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens (Abschnitt 4.2)
- Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen (Abschnitt 4.3)
- eine abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens, einschließlich der Angabe von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Abschnitt 4.4)
- ergänzende Informationen zur Begründung der vorgelegten Unterlagen (Abschnitt 4.5)

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

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

Ziel des Dossiers ist es, den medizinischen Nutzen und medizinischen Zusatznutzen von Sativex[®] Spray zur Anwendung in der Mundhöhle (Wirkstoffe: Delta-9-Tetrahydrocannabinol / Cannabidiol) in der Zusatzbehandlung von Patienten mit mittelschwerer bis schwerer Spastik aufgrund von Multipler Sklerose (MS), die nicht angemessen auf eine andere anti-spastische Arzneimitteltherapie angesprochen haben und die eine klinisch erhebliche Verbesserung von mit der Spastik verbundenen Symptomen während eines Anfangstherapieversuchs aufzeigen, darzustellen. Die Bewertung erfolgt anhand von RCT im direkten Vergleich zur zweckmäßigen Vergleichstherapie, die als die Fortführung der bisherigen pharmakologischen anti-spastischen Therapie definiert wird. Die Vortherapie ergibt sich aus der Summe der prozentualen Anteile aller bei den Patienten eingesetzten Arzneimittel (Mono- oder Kombinationstherapie), die für die vorliegende Indikation zugelassen sind. Die patientenrelevanten Endpunkte sind hierbei Morbidität (Verbesserung der Spastik; Globale Beurteilung der Veränderung hinsichtlich der Spastik durch Patienten, betreuende Personen und den Arzt) und Lebensqualität (Lebensqualität allgemein; Schlafqualität; Aktivitäten des täglichen Lebens; Gehfähigkeit) sowie Nebenwirkungen.

Datenquellen

Für die Bewertung des Nutzen und Zusatznutzens der Zusatztherapie mit Sativex[®] wurden alle in Bezug auf patientenrelevante Endpunkte aussagekräftigen RCT des pharmazeutischen Herstellers berücksichtigt sowie die, die in den Studienregistern und in der systematischen Literaturrecherche identifizierbar waren.

Ein-/Ausschlusskriterien für Studien

Eingeschlossen wurden alle RCT, die die Wirksamkeit der Wirkstoffkombination THC:CBD im Verhältnis 1:1 in der Zusatztherapie von mittelschwerer oder schwerer Spastik bei Patienten mit MS, die nicht angemessen auf eine andere anti-spastische Arzneimitteltherapie angesprochen haben und die eine klinisch erhebliche Verbesserung von mit der Spastik verbundenen Symptomen während eines Anfangstherapieversuchs aufzeigen, untersuchten. Diese RCT mussten Ergebnisse zu patientenrelevanten Endpunkten beschreiben und außerdem als Volltextpublikation zur Verfügung stehen.

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

Die Verzerrungsaspekte der Ergebnisse wurden für jede einzelne der eingeschlossenen Studien beschrieben und bewertet. Die Bewertung der einzelnen Aspekte erfolgte getrennt auf Ebene der einzelnen RCT sowie für jeden Endpunkt. Die Ergebnisse wurden in den im Dossier vorgegebenen Bewertungsbögen dokumentiert. Die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen aus den identifizierten RCT wurden für die Wirkstoffkombination bestehend aus THC und CBD im Verhältnis 1:1 im Vergleich zur zweckmäßigen Vergleichstherapie für jeden der patientenrelevanten Endpunkte tabellarisch

gegenübergestellt und ergänzend einzeln sowie zusammenfassend beschrieben. Abschließend wurde zur Beschreibung des Zusatznutzens sowie für dessen Ausmaß und Wahrscheinlichkeit eine Zusammenfassung und Bewertung der vorliegenden Ergebnisse als solche vorgenommen, unter Berücksichtigung der Validität der herangezogenen Endpunkte, der Studienqualität sowie der Evidenzstufe der vorliegenden Informationen.

Eine Meta-Analyse für die patientenrelevanten Endpunkte war für die Endpunkte Veränderung des Grads der Spastizität, gemessen auf der NRS und klinisch relevanter Rückgang der Schwere der Spastizität (Responder-Analyse), erhoben anhand der Anzahl der Patienten, für drei RCT möglich. Aufgrund des unterschiedlichen Designs der zugrundeliegenden RCT und der Heterogenität der Datenlage ließen sich für die übrigen patientenrelevanten Endpunkte keine Meta-Analysen anfertigen.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Der beträchtliche Zusatznutzen einer Zusatzbehandlung mit der Wirkstoffkombination THC:CBD 1:1 in der zugelassenen Indikation ergibt sich aus den angeführten Ergebnissen der RCT. Der Zusatznutzen der Zusatzbehandlung mit der Wirkstoffkombination THC:CBD 1:1 zeigt sich im Rückgang des Grads der Spastik, in einer höheren Rate der Patienten mit klinisch relevanter Verbesserung der Spastik, einer Reduktion der durchschnittlichen Anzahl an Spastiken pro Tag, einer Verbesserung der Schlafqualität, einer besseren globalen Beurteilung der Veränderung durch die Patienten, die betreuenden Personen und die behandelnden Ärzte, einer besser beurteilten Lebensqualität, einer Verbesserung der Aktivitäten des täglichen Lebens und der Gehfähigkeit. Die Häufigkeit von schwerwiegenden Nebenwirkungen mit möglichem Zusammenhang der Behandlung mit der Wirkstoffkombination THC:CBD 1:1 war gering.

Im direkten Vergleich zur Vergleichstherapie zeigt die Wirkstoffkombination THC:CBD 1:1 eine deutlichere Verbesserung der Spastik auf der 11-stufigen NRS und der VAS. Der Anteil der Patienten mit klinisch relevanter Verbesserung ist unter Zusatzbehandlung mit der Wirkstoffkombination THC:CBD 1:1 höher und war in der konfirmatorischen RCT statistisch signifikant. Die durchschnittliche Anzahl der Spastiken war in dieser RCT ebenfalls statistisch signifikant geringer unter der Wirkstoffkombination THC:CBD 1:1 gegenüber der Vergleichstherapie. Zudem ist die Schlafqualität stärker verbessert unter der Zusatzbehandlung mit der Wirkstoffkombination THC:CBD 1:1. Der Gesundheitszustand zeigt sich bei der Beurteilung auf einer visuellen Analogskala (VAS) im Rahmen der EQ-5D nur numerisch zugunsten der Wirkstoffkombination THC:CBD 1:1 verändert. Der Barthel Index als Maß für die Aktivitäten des täglichen Lebens bestätigt dieses Ergebnis.

Die Globalbeurteilung der klinischen Veränderung durch verschiedene Personen zeigt bei Patienten in der zugelassenen Indikationsstellung für die Wirkstoffkombination THC:CBD 1:1 (d.h. bei Patienten, die auf die Wirkstoffkombination THC:CBD 1:1 ansprechen) eine konsistente, statistisch signifikante Verbesserung unter der Zusatzbehandlung mit THC:CBD 1:1. Die verbesserte Bewertung der Veränderung zeigt sich in der Globalbeurteilung sowohl durch den Patienten selbst als auch durch den behandelnden Arzt und durch die betreuende

Person. Letztere bewerteten sowohl die Veränderung der funktionalen Möglichkeiten als auch die Veränderung der Transferierbarkeit des Patienten, die unter der Wirkstoffkombination THC:CBD 1:1 verbessert sind.

In allen Studien zeigte sich die Wirkstoffkombination THC:CBD 1:1 als gut verträglich.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Der Zusatznutzen einer Zusatztherapie mit Sativex[®] von Patienten mit mittelschwerer bis schwerer Spastik aufgrund von Multipler Sklerose (MS), die nicht angemessen auf eine andere anti-spastische Arzneimitteltherapie angesprochen haben, ist in dem hier vorliegenden Dossier belegt. Die im Dossier als Resultat der systematischen Übersicht dargelegten Nachweise zum Zusatznutzen einer Zusatztherapie mit der Wirkstoffkombination THC:CBD 1:1 berücksichtigen Ergebnisse von fünf verschiedenen RCT und entsprechen so den Anforderungen des Evidenzlevels 1. Dabei wurden vier der RCT an Patientenpopulationen durchgeführt, die der jetzt zugelassenen Anwendung von Sativex[®] in Deutschland entsprechen hinsichtlich einer unzureichenden Wirkung der Vortherapie der Spastik. Aus der fünften Studie wurde nur derjenige Teil der Patientenpopulation betrachtet, der den Einschlusskriterien dieses Nutzendossiers entsprach. Für diese Patienten lagen separate Ergebnisse zu den einzelnen Endpunkten vor, so dass eine separate Auswertung möglich war. Für die systematische Identifizierung der dem Dossier zugrunde liegenden Ergebnisse wurden öffentlich zugängliche Quellen zu klinischen Studien, Berichte zu klinischen Studien und Literaturdatenbanken wie im Dossier dargestellt, sowie die beim pharmazeutischen Hersteller verfügbaren Informationen zu relevanten klinischen Studien voll berücksichtigt. Insofern liegt Ergebnisvollständigkeit vor.

Das Ausmaß des Zusatznutzens der Zusatztherapie mit Sativex[®], im Anwendungsgebiet, ist beträchtlich. Unter Zusatztherapie mit Sativex[®] wird für Patienten mit mittelschwerer bis schwerer Spastik aufgrund von Multipler Sklerose (MS), die nicht angemessen auf eine andere anti-spastische Arzneimitteltherapie angesprochen haben und die eine klinisch erhebliche Verbesserung von mit der Spastik verbundenen Symptomen während eines Anfangstherapieversuchs aufzeigen, ein bisher nicht erreichter patientenrelevanter therapeutischer Zusatzeffekt erzielt. Unter der Wirkstoffkombination THC:CBD 1:1 wird eine gegenüber der zweckmäßigen Vergleichstherapie deutliche Verbesserung des allgemeinen Zustandes der Patienten erreicht. Dies zeigt sich in der Verbesserung des Grads der Spastizität, gemessen auf der NRS sowie in der höheren Anzahl der Patienten, die eine klinisch relevante Verbesserung der Schwere der Spastik aufzeigen sowie durch die deutliche Abnahme der durchschnittlichen Anzahl der Spastiken pro Tag. Mehr Patienten nehmen unter Zusatztherapie mit der Wirkstoffkombination THC:CBD 1:1 insgesamt eine spürbare Linderung der Symptomatik wahr, was sowohl von den Patienten selbst als auch durch die den Patienten betreuenden oder behandelnden Personen berichtet wird. Ein weiterer funktioneller Parameter, der Barthel-Index, der die Aktivitäten des täglichen Lebens misst, zeigt ebenfalls eine erhebliche Verbesserung unter der Wirkstoffkombination THC:CBD 1:1.

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.

Ziel des Dossiers ist es, den medizinischen Nutzen und medizinischen Zusatznutzen von Sativex® Spray zur Anwendung in der Mundhöhle (Wirkstoffe: Delta-9-Tetrahydrocannabinol/ Cannabidiol) in der Zusatzbehandlung von Patienten mit mittelschwerer bis schwerer Spastik aufgrund von Multipler Sklerose (MS), die nicht angemessen auf eine andere anti-spastische Arzneimitteltherapie angesprochen haben und die eine klinisch erhebliche Verbesserung von mit der Spastik verbundenen Symptomen während eines Anfangstherapieversuchs aufzeigen, darzustellen. Die Bewertung erfolgt anhand von RCT im direkten Vergleich zur zweckmäßigen Vergleichstherapie, die als die Fortführung der bisherigen pharmakologischen anti-spastischen Therapie definiert wird. Die Vortherapie ergibt sich aus der Summe der prozentualen Anteile aller bei den Patienten eingesetzten Arzneimittel (Mono- oder Kombinationstherapie), die für die vorliegende Indikation zugelassen sind. Die patientenrelevanten Endpunkte sind hierbei Morbidität (Verbesserung des Schweregrads der Spastik; Anzahl der Patienten mit einer klinisch relevanten Verbesserung der Schwere der Spastik; Anzahl der Spastiken pro Tag; Globale Beurteilung der Veränderung hinsichtlich der Spastik durch Patienten, betreuende Personen und den Arzt) und Lebensqualität (allgemeine Lebensqualität, Gesundheitszustand, Schlafqualität, Aktivitäten des täglichen Lebens, Gehfähigkeit).

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.

Die Nutzenbewertung basiert auf geeigneten randomisierten, klinischen Studien und basiert so auf Informationen auf dem höchsten Evidenzlevel. Es wurde die Wirkstoffkombination THC:CBD im Verhältnis 1:1 als Zusatzbehandlung (Add-on Therapie) unter Fortführung der Vortherapie der Spastik direkt mit der Fortführung der Vortherapie plus Placebo bei Patienten mit mittelschwerer bis schwerer Spastik aufgrund von MS verglichen, die nicht angemessen auf eine andere anti-spastische Arzneimitteltherapie angesprochen haben. Als zweckmäßige Vergleichstherapie wurde die Fortführung der bisherigen pharmakologischen anti-spastischen Therapie angesehen. Die Vortherapie ergibt sich aus der Summe der prozentualen Anteile aller bei den Patienten eingesetzten Arzneimittel, die für die vorliegende Indikation zugelassen sind (siehe Modul 3A). Da ein Behandlungseffekt bezüglich der Spastik bereits sehr kurzfristig einsetzt, wurde vorab bezüglich geeigneter Studien keine Mindestdauer für den Behandlungszeitraum definiert.

Primär wurden, auf Basis der vorliegenden Zulassungsstudien, folgende Endpunkte identifiziert:

1. Ermittlung der Spastizität (NRS, Ashworth Skala)
2. Ermittlung des Grads der Behinderung (Expanded Disability Status Scale [EDSS])
3. Häufigkeit der Spasmen (Anzahl der Spastiken pro Tag)
4. Gehfähigkeit (10 m-Gehstreckentest)
5. Schlafqualität (NRS)
6. Aktivitäten des täglichen Lebens (Barthel-Index)
7. Gesundheitszustand (EQ-5D)
8. Lebensqualität (VAS QoL, SF-36 und MSQoL)
9. Globale Beurteilung der Spastik (PGIC, CGIC, SGIC)
10. Depressionen (Beck Depression Inventory)

11. Nebenwirkungen (GCP-konforme Dokumentation)
12. Klinische Analysen (Blut- und Urinanalyse, Vitalzeichen, EKG)
13. Grad der Berauschtigkeit (NRS)

Von diesen in den klinischen Studien identifizierten Endpunkten wurden die Endpunkte Depression und Grad der Berauschtigkeit durch die Dokumentation der Nebenwirkungen abgedeckt. Sie sind für die zugrundeliegende Erkrankung nicht relevant. Der EDSS wird hier nicht weiter betrachtet, da dieser Score den Parameter Spastik weder direkt abbildet noch indirekt zu messen in der Lage ist. Folgende Endpunkte wurden als patientenrelevant eingestuft und für die vergleichende Betrachtung des Zusatznutzens herangezogen (Tabelle 4-1):

Tabelle 4-1: Patientenrelevante Endpunkte

Definierte Endpunkte laut Zulassungsstudien	Patientenrelevante Endpunkte nach § 3 Verfo des G-BA
Verbesserung der Spastik: Schweregrad der Spastik und Häufigkeit der Spasmen pro Tag	Verbesserung des Gesundheitszustandes (Morbidität)
Globale Beurteilung der Veränderung hinsichtlich der Spastik durch den Patienten (SGIC), die betreuende Person (CGIC) und den Arzt (PhGIC)	
Gesundheitszustand (EQ-5D)	
Einfluß auf die Lebensqualität (VAS QoL, SF-36 und MSQoL)	Verbesserung der Lebensqualität
Schlafqualität des Patienten	
Aktivitäten des täglichen Lebens	
Gehfähigkeit	
Auftreten von Nebenwirkungen	Verringerung der Nebenwirkungen

Der patientenrelevante Endpunkt Mortalität (Verlängerung des Überlebens) wurde nicht primär in die Nutzenbewertung einbezogen, da Sativex[®] die Symptome des Krankheitsbildes mildert, aber keinen Einfluss auf den Krankheitsverlauf hat. Der Endpunkt Mortalität wird hier nicht für sich abgebildet, findet sich allerdings im Endpunkt Nebenwirkungen, speziell in den schwerwiegenden Nebenwirkungen, wieder.

Sämtliche Kriterien, die für den Einschluss von Studien in die vorliegende Bewertung des Zusatznutzens angewandt wurden, sind in nachfolgender Abbildung zusammengefasst.

Einschlusskriterien	
E1	Die Studie ist ein RCT.
E2	Die Studie präsentiert Daten zur Wirkstoffkombination THC:CBD im Verhältnis 1:1.
E3	Die Studie präsentiert Daten zur Patientenpopulation „Multiple Sklerose“.
E4	Die Studie präsentiert Daten zum Anwendungsgebiet „mittelschwere bis schwere Spastik“.
E5	Die Studie präsentiert Daten im Vergleich zur Fortführung der bisherigen Therapie plus Placebo (Vergleichstherapie).
E6	Die Studie beinhaltet mindestens einen der in 4.2.2 benannten patientenrelevanten Endpunkte.
E7	Volltextpublikation verfügbar.

Abbildung 1: Einschlusskriterien für die Berücksichtigung von Studien in der Bewertung des Zusatznutzens.

Die Einschlusskriterien gelten als erfüllt, wenn in der zu bewertenden Studie mindestens 80 % der Patienten die Einschlusskriterien erfüllen.

Die zweckmäßige Vergleichstherapie wird definiert als die Fortführung der bisherigen pharmakologischen anti-spastischen Therapie. Aus diesem Grund erfolgt ein direkter Vergleich der Wirkstoffkombination THC:CBD 1:1 als Add-on Therapie zur bisherigen pharmakologischen anti-spastischen Therapie gegenüber der bisherigen pharmakologischen anti-spastischen Therapie plus Placebo.

Es wurden keine zusätzlichen Ausschlusskriterien definiert.

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 bibliographische Literatursuche wurde systematisch in folgenden Literaturdatenbanken durchgeführt: Medline, EMBASE und Cochrane. Die Suche wurde für das zu bewertende Arzneimittel und die Vergleichstherapie gemeinsam durchgeführt, da in den nachfolgenden Kapiteln die direkte Darstellung des Zusatznutzens anhand des direkten Vergleichs zur zweckmäßigen Vergleichstherapie erfolgt.

In der Datenbank EMBASE wurde die Einschränkung getroffen, nur in Embase und nicht noch einmal in Medline zu suchen, um im Vorfeld eventuelle Duplikate auszuschließen. Weitere Einschränkungen wurden nicht vorgenommen.

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.

Folgende Studienregister wurden zur Identifizierung relevanter klinischer Studien durchsucht: clinicaltrials.gov, clinicalstudyresults.org, Clinical Trials Registry Platform Search Portal (ICTRP Search Portal) der WHO und die EudraCT Datenbank der EMA.

Es 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 Auswahl der geeigneten Studien aus den identifizierten Studien erfolgte durch zwei Personen unabhängig voneinander. Abweichende Beurteilungen wurden im abschließenden Konsensverfahren durch beide Personen geklärt und entsprechend dokumentiert.

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.

Die Verzerrungsaspekte der Ergebnisse wurden für jede eingeschlossene Studie beschrieben. Die Bewertung der einzelnen Aspekte wurde für jeden Endpunkt im Bewertungsbogen laut Anhang 4-G dokumentiert.

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.

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, 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.

In Modul 4 werden zur Nutzenbewertung ausschließlich randomisierte klinische Studien berücksichtigt. Die Darstellung der Studien entsprechen dabei mindestens den in diesem Dossier zitierten CONSORT Anforderungen Item 2b bis 14.

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. 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.

Die Betrachtung des Zusatznutzens von Sativex® bezieht sich auf die Zusatztherapie von Patienten mit mittlerer bis schwerer Spastik bei MS, die unter ihrer bisherigen Therapie der Spastik nur unzureichend eingestellt sind. An den zur Nutzenbewertung herangezogenen Studien nahmen mehr Frauen (60 %) als Männer (40 %) teil. Der Altersdurchschnitt der gesamten Patientenpopulation betrug ca. 49 Jahre.

Zur Darstellung des Zusatznutzens werden folgende patientenrelevante Endpunkte betrachtet:

- Verbesserung des Grads der Spastik und Anzahl der Patienten mit einem klinisch relevanten Rückgang der Schwere der Spastik:

Die Spastik wurde auf der 11-stufigen NRS (0-10) gemessen, die in vielen Indikationen zur Messung von Endpunkten eingesetzt wird (1). Die NRS ist geeignet, um in klinischen Studien relevante Änderungen der Spastik abzubilden (2). Die 11-stufige NRS korreliert statistisch signifikant mit dem Grad der Spastik sowie der globalen Beurteilung der klinischen Veränderung (2). Die individuelle Verbesserung des NRS Scores für den Spastizitätsgrad muss mindestens 18 % betragen, damit von einer klinischen Relevanz der globalen Verbesserung gesprochen werden kann. Eine

Studie benutzte statt der NRS die 0-100 mm VAS, die für die Meta-Analyse in die 11-stufige NRS linear umgerechnet wurde (3, 4)

Eine Verbesserung von 30 % (entspricht 3 Punkten) oder mehr auf der 11-stufigen NRS bedeutet eine deutliche klinische Verbesserung der Spastik für den Patienten (2), deutlich über das Mindestmaß hinaus, das als klinisch relevant angesehen wird.

Der patientenrelevante Endpunkt „Anzahl der Patienten“ beschreibt eine Verbesserung des Grades der Spastik um mindestens 30 % auf der 11-stufigen NRS zwischen den Behandlungsgruppen. Darüber hinaus wurde untersucht, wie viele Patienten sich im Vergleich zum Ausgangswert vor Beginn der Therapie mit der Wirkstoffkombination THC:CBD 1:1 in ihrer Spastizität um mehr als 30 % auf der NRS verbesserten. Ergänzend wurde zur Sensitivitätsanalyse der Endpunkt auf Basis einer Verbesserung um mindestens 50 % gegenüber dem Zustand vor der Behandlung mit THC:CBD 1:1 auf der 11-stufigen NRS analysiert.

- Die Häufigkeit von Spastiken pro Tag:

In diesem Endpunkt wird die Anzahl von Spastiken pro Tag berichtet. Jede Verminderung der Anzahl einschließender Spastiken ist patientenrelevant, da jeweils eine akute Verschlechterung des Gesundheitszustands vermieden wird.

- Die globale Beurteilung der Veränderung hinsichtlich der Spastik durch den Patienten (SGIC), die betreuende Person (CGIC) und den Arzt (PhGIC):

In der klinischen Praxis werden Beurteilungen durch den Patienten oder die betreuende Person häufig zur klinischen Beurteilung einer Spastik durch den Arzt mit herangezogen (5-7). Die Erfassung der Globalbeurteilung in den RCT erfolgt auf einer kategoriellen 7-stufigen VRS (8). Die Einschätzung des klinischen Zustands bzw. dessen Veränderung auf einer 7-stufigen VRS stellt eine Methode zur standardisierten Beurteilung von Erkrankungsbildern dar, deren Anwendung auch bei MS-Patienten validiert ist (2). Die Erfassung auf der 7-stufigen VRS ist weniger sensitiv im Vergleich zu anderen Methoden und die Veränderung um eine Kategorie bildet bereits einen klinisch relevanten Unterschied ab (9, 10).

- Gesundheitszustand des Patienten:

Die Veränderung des Gesundheitszustandes ist als patientenrelevanter Parameter explizit im § 3 der VerfO genannt und wird mittels des validierten EQ-5D Fragebogens erhoben (11). Als Endpunkte wurden die beiden unabhängigen Parameter des EQ-5D Instruments, der EQ-5D Utility Score und die Beurteilung des Gesundheitszustands auf der VAS, betrachtet. Der EQ-5D ist in der Anwendung bei MS Patienten validiert (12).

- Die Lebensqualität des Patienten:

Die Lebensqualität ist ein patientenrelevanter Parameter und wird mittels des MSQoL-54, des SF-36 und des VAS QoL erhoben. Der SF-36 ist eines der meistverwendeten Instrumente zur Bestimmung der allgemeinen Lebensqualität und unterscheidet zwischen Patienten verschiedener Schweregrade derselben Krankheit. Der MSQoL-54 ist ein multidimensionales Messinstrument zur Bestimmung der gesundheitsbezogenen Lebensqualität, das allgemeine und MS-spezifische Elemente in einem Instrument zusammenfasst (13, 14). Die Entwickler dieser Methode benutzen dabei den SF-36 als allgemeinen Bestandteil und erweiterten diesen um 18 MS-spezifische Elemente wie Erschöpfung oder kognitive Funktionen. Das MSQoL ist ein strukturierter Fragebogen, der von den Patienten generell selbst ausgefüllt werden kann. Eine Bearbeitung des Fragebogens durch einen Interviewenden gerade bei Patienten mit Einschränkungen im Bereich der oberen Extremitäten oder des visuellen Bereichs ist auch möglich. Es gibt für den MSQoL-54 keinen einzelnen Wert für alle Bewertungspunkte zusammen, sondern zwei zusammenfassende Werte (physische und mentale Gesundheit) sowie 12 Unterskalen (14). In einer Studie wurde die Lebensqualität auf einer 0-100 mm VAS (VAS QoL) von den Patienten bestimmt.

- Die Schlafqualität des Patienten:

Die Schlafqualität als Endpunkt wurde separat auf einer 11-stufigen NRS erfasst, da Schlaf ein anerkannt wichtiger Faktor der Gesundheit ist, mit unmittelbarem Einfluss auf den patientenrelevanten Endpunkt Lebensqualität (15, 16). Bei Patienten mit MS kann Schlafunterbrechung zusätzlich die Symptome der Fatigue über den Tag verstärken (17). In der Literatur wird kein klinisch relevanter Mindestwert beschrieben.

- Aktivitäten des täglichen Lebens:

Die Fähigkeit zur Ausübung von Aktivitäten des täglichen Lebens ist wichtig für die individuelle Selbstständigkeit im Alltag. Der Endpunkt Barthel Index (BI) ist ein Instrument zur quantitativen Erfassung der Fähigkeiten zur Ausübung der Aktivitäten des täglichen Lebens. Der BI ist bereits bei über 50 Studien mit MS Patienten eingesetzt worden (18). Der BI liefert eine valide Messung der Einschränkungen der Aktivitäten des täglichen Lebens. Es werden 10 Elemente gemessen (zwei werden auf einer Skala von 0 oder 5, sechs auf einer Skala von 0, 5 oder 10 und zwei auf einer Skala von 0,5, 10 oder 15), die zusammen einen Wert von 100 ergeben. Die minimale klinisch relevante Veränderung des BI liegt bei 1,85 Punkten (19).

- Gehfähigkeit (10 m-Gehstrecke):

Der Endpunkt Gehfähigkeit wird mittels des 10 m-Gehstreckentests bestimmt. Dieser misst die Zeit, die für das Zurücklegen einer Gehstrecke von 10 Metern benötigt wird (20). Der Test wird als ein Indikator für die auch vom Grad der Spastik abhängige Mobilität angesehen. Der Gehstreckentest wird auch bei MS Patienten angewendet (21). Bei einer Veränderung der Gehstrecken-Geschwindigkeit von mindestens

0,05 m/s ist von einer klinisch relevanten Verbesserung der Gehfähigkeit auszugehen (22).

- Nebenwirkungen

Ein weiterer patientenrelevanter Endpunkt ist die Verringerung der Nebenwirkungen. Es wird das Auftreten von Nebenwirkungen und insbesondere der schwerwiegenden Nebenwirkungen betrachtet, da bei derartigen Nebenwirkungen mit höherer Wahrscheinlichkeit von einem hohen Schadenspotential für den Patienten ausgegangen werden kann.

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 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

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

⁵Deeks 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.

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

Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

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

Es wurde eine Meta-Analyse mit Ergebnissen aus den ITT-Populationen von drei RCT zu den Endpunkten Veränderung des Grads der Spastizität und klinisch relevanter Rückgang der Schwere der Spastizität (Responder, mind. 30 %), gemessen auf der NRS, angefertigt. Für alle übrigen patientenrelevanten Endpunkte war die Durchführung einer Meta-Analyse aus Gründen der Heterogenität nicht angebracht. Die Meta-Analyse der o. g. Endpunkte berücksichtigte die Rohdaten aus den ITT-Populationen der drei Studien und betrachtet sowohl die Mittelwertdifferenz im NRS Score als auch die Odds ratio für einen klinisch relevanten Rückgang des Grads der Spastik auf Basis von Modellen mit zufälligen Effekten. Ist die Heterogenität nicht bedeutsam ($< 20\%$), wird der gemeinsame gepoolte Effekt inklusive Konfidenzintervall dargestellt. Bei bedeutsamer Heterogenität werden die Ergebnisse nur in begründeten Ausnahmefällen gepoolt und gegebenenfalls einer Sensitivitätsanalyse unterzogen. Außerdem wird untersucht, welche Faktoren diese Heterogenität möglicherweise erklären könnten.

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.

In der Literatur ist für die Verbesserung des Grads der Spastizität auf der 11-stufigen NRS eine klinisch relevante Veränderung von 18 % genannt (2). Im Rahmen der Responder-Analyse wurde ein primärer Cut-off von 30 % der Verbesserung auf der NRS festgelegt. Dieser Wert wurde gewählt, um mit einem sicheren Maß von einer Verbesserung ausgehen zu können. Im Rahmen der Sensitivitätsanalyse des Endpunktes Anzahl der Patienten mit einem

klinisch relevanten Rückgang der Schwere der Spastizität wurde auch die Verbesserung des Grads der Spastizität um 50 % untersucht.

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.

NA - Es wurden keine relevanten Subgruppen analysiert, da in den Studien nicht nach Geschlecht, Alter und Krankheitsschwere unterschieden wurde und auch keine weiteren Subgruppenanalysen a priori definiert wurden.

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 Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist¹¹. (23-26)Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen¹².

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

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

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

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

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

¹²Song 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.

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

- 1) 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.*
- 2) Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*
- 3) Art der Prüfung der Konsistenz zwischen den Ergebnissen direkter und indirekter Vergleiche.*
- 4) 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).*
- 5) Art und Umfang von Sensitivitätsanalysen.*

NA – Die im Rahmen des Dossiers vorgelegte Evidenz zeigt den Zusatznutzen der Zusatztherapie mit Sativex® im direkten Vergleich zur zweckmäßigen Vergleichstherapie.

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 4-2: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studien-dauer	Therapiearme
GWCADD01 A safety and tolerability study of medicinal cannabis extracts for chronic refractory spasticity and neurogenic pain	Nein	abgeschlossen	11 Wochen 08/2001 - 05/2002	GW-1000/1-02: THC/CBD, extract of Cannabis sativa L. GW-2000-02 THC extract of Cannabis sativa L. GW-3000-02 CBD extract of Cannabis sativa L. Matching placebo:GW-4001-01
GWN 10005 An “N of 1” placebo controlled study of cannabis based medicine extracts in treatment of multiple sclerosis, spinal cord injury, peripheral nerve injury or rheumatoid arthritis.	Nein	abgebrochen	12 Wochen 10/2001– 11/2001	GW-1000/1-02: Δ9-tetrahydro- cannabinol THC/CBD extract of Cannabis sativa L. GW-2000-02 THC extract of Cannabis sativa L. GW-3000-02 CBD extract of Cannabis sativa L. Matching placebo: GW-4001-01
GWN19902 A preliminary double-blind, randomised crossover, placebo-controlled investigation of the therapeutic profile of three cannabis-based medicine extracts in patients with multiple sclerosis, spinal cord conditions, peripheral nerve injury or central nervous system damage associated with	Nein	abgeschlossen	10 Wochen 07/2000 - 11/2001	GW-1000/1/2 THC/CBD extract of Cannabis sativa L. GW-2000 THC extract of Cannabis sativa L. GW-3000 CBD extract of Cannabis sativa L.

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studien-dauer	Therapiearme
vascular, traumatic, infective, genetic, neoplastic and metabolic disease; any of which has resulted in pain and/or spasticity ⁷				Placebo
GWN19904 A preliminary placebo-controlled investigation of the therapeutic profile of three cannabis based medicine extracts in patients with multiple sclerosis and other neurological conditions resulting in pain and/or spasticity, or arthritic conditions	Nein	abgeschlossen	8 Wochen 09/2000-03/2002	GW-1000/1 THC/CBD extract of Cannabis sativa L. GW-2000/1 THC extract of Cannabis sativa L. GW-3001/2 CBD extract of Cannabis sativa L. GW-4000/1: Placebo
GWMS0001 A double blind, randomised, parallel group, placebo-controlled trial of a combination of delta-9-tetrahydrocannabinol (THC) and cannabidiol (CBD) in patients with multiple sclerosis, followed by an open label assessment and study extension.	Nein	abgeschlossen	6Wochen 05/2001 - 07/2002	GW-1000-02 THC/CBD extract of Cannabis sativa L Placebo
GWMS0106 A Double Blind, Randomised, Parallel Group Study to Assess the Efficacy, Safety and Tolerability of Cannabis Based Medicine 1:1 THC/CBD Compared With Placebo for the Treatment of Spasticity in Patients With Multiple Sclerosis.	Ja	abgeschlossen	8 Wochen (2 Wochen Baseline + 6 Wochen treatment) 06/2002 – 03/2004	THC/CBD (Sativex®) Placebo
GWSP0604 A two-phase, phase 3 study of the safety and efficacy of Sativex®, in the symptomatic relief of spasticity in subjects with spasticity due to multiple sclerosis: PHASE A – Single-blind response assessment; PHASE B – Double-blind randomized, placebo controlled, parallel group study	Ja	abgeschlossen	19 Wochen 01/2008-01/2009	THC/CBD (Sativex®) Placebo
GWSP0702 A Placebo Controlled, Parallel Group, Randomised Withdrawal	Ja	abgeschlossen	5 Wochen 11/2007-	THC/CBD (Sativex®)

Studienbezeichnung	Zulassungs- studie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studien- dauer	Therapiearme
Study of Subjects With Symptoms of Spasticity Due to Multiple Sclerosis Who Are Receiving Long-term GW-1000-02 (Sativex®).			01/2009	Placebo
GWCL0403 A double blind, randomised, placebo controlled, parallel group study of Sativex®, in subjects with symptoms of spasticity due to multiple sclerosis	Ja	abgeschlossen	14 Wochen 03/2005- 12/2005	THC/CBD (Sativex®) Placebo

Geben Sie an, welchen Stand die Information in Tabelle 4-2 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 wurden am 01. November 2011 festgestellt.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-2 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-3: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
GWCADD01 A safety and tolerability study of medicinal cannabis extracts for chronic refractory spasticity and neurogenic pain	Das Einschlusskriterium E3 ist nicht erfüllt. Nur 2 von 8 Patienten sind MS-Patienten.
GWN 10005 An “N of 1” placebo controlled study of cannabis based medicine extracts in treatment of multiple sclerosis, spinal cord injury, peripheral nerve injury or rheumatoid arthritis.	Das Einschlusskriterium E3 ist nicht erfüllt. Es handelte sich nicht um einen MS-Patienten.
GWN19902 A preliminary double-blind, randomised crossover, placebo-controlled investigation of the therapeutic profile of three cannabis-based medicine extracts in patients with multiple sclerosis, spinal cord conditions, peripheral nerve injury or central nervous system damage associated with vascular, traumatic, infective, genetic, neoplastic and metabolic disease; any of which has resulted in pain and/or spasticity ⁷	Das Einschlusskriterium E4 ist nicht erfüllt. Zu den 17 MS-Patienten von insgesamt 25 Patienten liegt keine Unterscheidung zur Stärke der Spastizität in leicht, mittelschwer und schwer vor. Außerdem erfolgte keine separate Auswertung der Ergebnisse für diese Patienten.
GWN19904 A preliminary placebo-controlled investigation of the therapeutic profile of three cannabis based medicine extracts in patients with multiple sclerosis and other neurological conditions resulting in pain and/or spasticity, or arthritic conditions	Das Einschlusskriterium E4 ist nicht erfüllt. Zu den 14 MS-Patienten von insgesamt 27 Patienten liegt keine Unterscheidung zur Stärke der Spastizität in leicht, mittelschwer und schwer vor.
GWSP0702 A Placebo Controlled, Parallel Group, Randomised Withdrawal Study of Subjects With Symptoms of Spasticity Due to Multiple Sclerosis Who Are Receiving Long-term GW-1000-02 (Sativex [®]).	Das Einschlusskriterium E4 ist nicht erfüllt. Es ist unklar, ob die Patienten vor der Erstgabe von Sativex [®] mittelschwere bis schwere Spastik oder nur leichte Spastik hatten.

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, Abstracts 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.

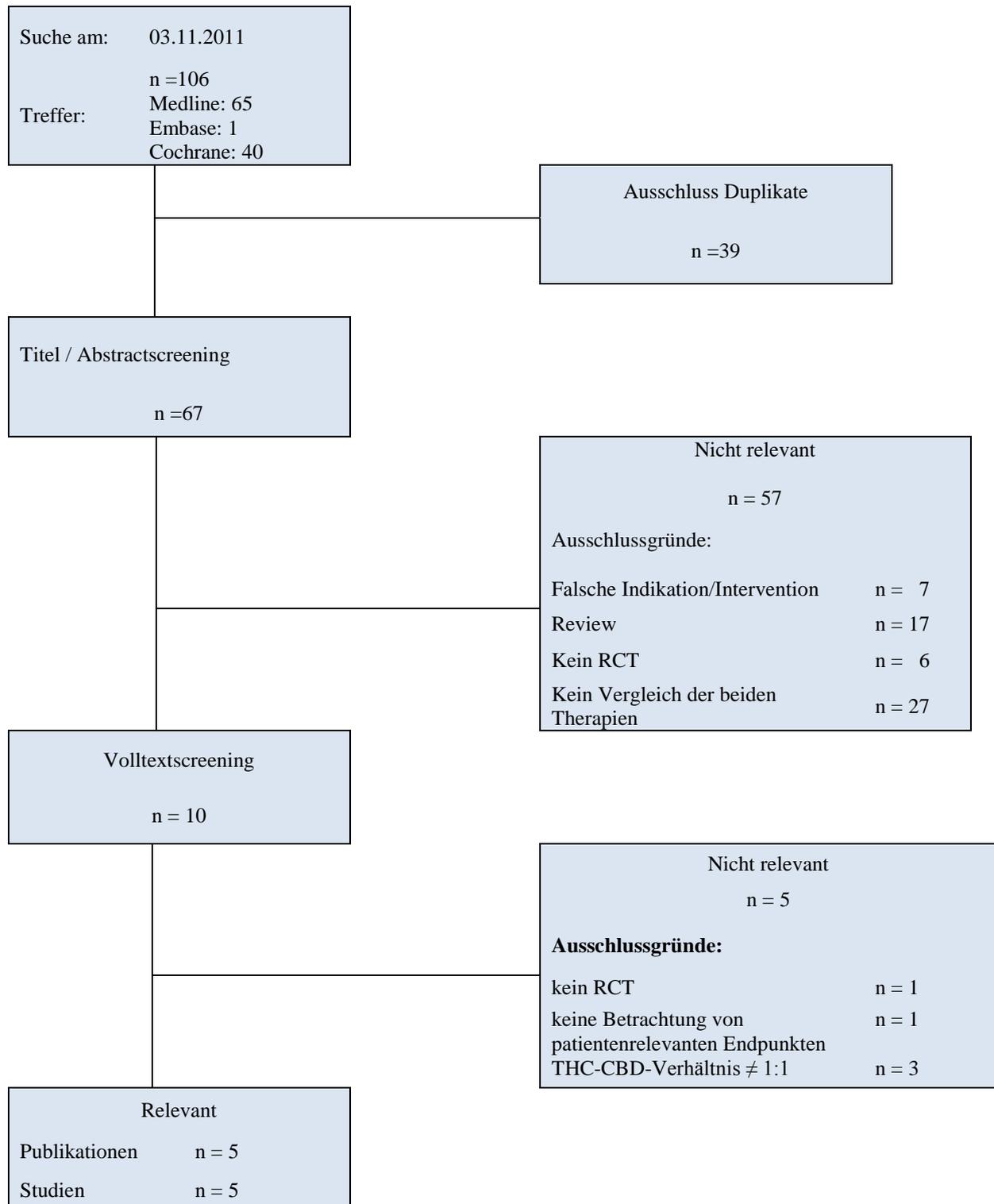


Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

4.3.1.1.3 Studien aus der Suche in Studienregistern

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-2) 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 4-4: 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)
GWMS0106	clinicaltrials.gov [NCT00711646], WHO-ICTRP [NCT00711646](27)	Ja	Ja
GWCL0403	clinicaltrialsregister.eu [EudraCT Nummer: 2004-002509-63](28)	Ja	Ja
GWSP0604	clinicaltrials.gov [NCT00681538], WHO-ICTRP [NCT00681538](29) clinicaltrialsregister.eu [EudraCT Nummer: 2006-005910-11]	Ja	Ja
NEU-CAN-04	clinicaltrials.gov [NCT0020243] (30)	Nein	Nein

a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.
b: n. d. = bibliografische Literaturrecherche nicht durchgeführt.

Bei der Studie NEU-CAN-04 handelt es sich um eine laufende Studie, aus der noch keine Ergebnisse vorliegen.

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 Vergleichs-

therapie. 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 4-5: 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])
placebokontrolliert						
GWMS0001	Nein	Ja	Nein	Ja(4)	Nein	Ja (31)
GWMS0106	Ja	Ja	Nein	Ja (32)	Ja (27)	Ja (33)
GWCL0403	Ja	Ja	Nein	Ja (34)	Ja(28)	Ja (35)
GWSP0604	Ja	Ja	Nein	Ja (36)	Ja (29)	Ja (37)
ARA2009	Nein	Nein	Ja	Nein	Nein	Ja (38)
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
Es liegen keine aktiv kontrollierten RCT vor.						
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.						

Es wurden insgesamt fünf Studien identifiziert, die für die Nutzenbewertung für Sativex[®] Spray zur Anwendung in der Mundhöhle als Add-On Therapie für MS-Patienten mit mittelschwerer bis schwerer Spastik, die bisher nicht angemessen auf eine andere anti-spastische Arzneimitteltherapie angesprochen haben, in Frage kommen.

In der Studie GWMS0001 nahmen insgesamt 160 MS-Patienten teil, von denen 140 Symptome der Spastik aufwiesen. Allerdings hatten insgesamt nur 37 Patienten als primäres Symptom die Spastik mit einem VAS-Score von mehr als 50 auf der 0-100 mm Skala. Da die Ergebnisse dieser Population separat beschrieben wurden, ist ein Einschluss dieser Studie in den Studienpool gerechtfertigt.

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 4-6: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
GWMS0001	Multizentrisch, randomisiert, doppelblind, Placebo kontrolliert	Erwachsene Patienten mit MS und vorausgesetztem Schweregrad in einer oder mehreren der folgenden Kategorien: Schmerz, Spastizität, Muskelspasmen, Blasenproblemen oder Tremor	GW-1000-02 oder GW-1001- 02 Add-on (n= 80) GW-4001-01 (Placebo) (n= 80)	Screening: 1-2 Wochen 6-wöchige aktive Behandlungsphase (placebo-kontrolliert) 4-wöchige open label- Phase Danach Möglichkeit der Teilnahme an einer Langzeitstudie	UK 05/2001- 07/2002	Veränderung der Spastizität mit Hilfe der VAS (visuelle Analogskala) oder MAS (modifizierte Ashworth Skala); Schlafqualität; Alltagsfunktion (Barthel-Index); 10-Meter-Gehtest; Auftreten von Nebenwirkungen
GWMS0106	Multizentrisch, multinational, randomisiert, doppelblind, Placebo kontrolliert	Erwachsene Patienten mit Spastik bedingt durch MS in zwei Muskelgruppen und nicht zufriedenstellend eingestellter Spastik.	Sativex® Add- on (n= 124) Placebo Add-on (n= 65)	Screening: 2 Wochen 6-wöchige aktive Behandlungsphase Nachbeobachtung: 7- 10 Tage	UK und Rumänien 06/2002 – 03/2004	Veränderung des Schweregrads der Spastik beurteilt auf Numerischer 11-Punkteskala (NRS); Globale Gesamteinschätzung der Veränderungen durch den Patienten (SGIC), Auftreten von Nebenwirkungen
GWCL0403	Multizentrisch, multinational, randomisiert, doppelblind, Placebo kontrolliert	Erwachsene Patienten mit Spastik bedingt durch MS und nicht zufriedenstellend eingestellter Spastik.	Sativex® Add-on (n= 167) Placebo Add-on (n= 170)	Screening: 1 Woche 14-wöchige aktive Behandlungsphase Nachbeobachtung: 28 Tage	14 Zentren in UK und 8 Zentren in Tschechien. 03/2005 - 12/2005	Veränderung des Schweregrads der Spastik beurteilt auf Numerischer 11-Punkteskala (NRS); Globale Gesamteinschätzung der Veränderungen durch die Bezugspersonen (CGIC), Gesundheitszustand (EQ-5D), Lebensqualität (MSQoL-54), Schlafqualität Alltagsfunktionen(Barthel-Index),

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

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
						10-Meter-Gehtest, Auftreten von Nebenwirkungen
GWSP0604	<p>Multizentrische, multinationale Studie in 2 Phasen</p> <p>Phase A: Einfach verblindet</p> <p>Phase B: Doppelblind, randomisiert, Placebo kontrolliert</p>	<p>Erwachsene; MS mit nicht vollständig eingestellter Spastik unter anti-spastischer Medikation oder Therapieversuch.</p> <p>Einarmig Sativex® über 4 Wochen, Responder (47 %, n=241) randomisiert für Sativex® oder Placebo über 12 Wochen. Die laufende Spastik-Medikation der Studienteilnehmer wurde fortgesetzt</p>	<p>Phase A (nicht randomisiert) Sativex® n=572</p> <p>Phase B: (randomisiert) n=241</p> <p>Sativex® Add-on (n=124)</p> <p>Placebo Add-on (n=117)</p>	<p>Phase A: Run-in: 4- wöchige aktive Behandlungsphase (alle)</p> <p>Phase B: 12-wöchige randomisierte Behandlungsphase (nur Responder aus Phase A)</p> <p>Follow-up 14 Tage nach Studienende</p>	<p>52 Studienzentren (18 Zentren in UK, 11 Zentren in Spanien, 10 Zentren in Polen, 8 Zentren in Tschechien 5 Zentren in Italien)</p> <p>01/2008 – 01/2009</p>	<p>Veränderung des Schweregrads der Spastik beurteilt auf Numerischer 11-Punkteskala (NRS);</p> <p>Häufigkeit der Spastiken,</p> <p>Globale Gesamteinschätzung der Veränderungen durch den Patienten (SGIC),</p> <p>Gesamteinschätzung der Veränderungen durch die Bezugspersonen (CGIC)</p> <p>Globale Gesamteinschätzung der Veränderungen durch den Investigator (PhGIC),</p> <p>Gesundheitszustand (EQ-5D),</p> <p>Lebensqualität (SF-36),</p> <p>Schlafqualität</p> <p>Alltagsfunktionen (Barthel-Index),</p> <p>Modified Ashworth Scale,</p> <p>10-Meter-Gehtest,</p> <p>Auftreten von Nebenwirkungen</p>
ARA2009	monozentrisch,	Erwachsene Patienten	n=17	8 Wochen	Keine Angaben	Pathophysiologische Symptome,

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

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
	Placebo kontrolliert, parallel, randomisiert, cross-over	mit Spastik bedingt durch MS in zwei Muskelgruppen und unzureichendem Ansprechen auf eine Vorbehandlung	Sativex® Add-on (n=17) Placebo Add-on (n=17)	3-wöchige Behandlungsphase, dann 2-wöchige Wash-out Periode, dann 3-wöchige Behandlungsphase (cross-over Behandlung)	zu Zentren 2004	Kognitive Performance, Lebensqualität und körperliche und psychische Verfassung, Auftreten von Nebenwirkungen

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

Studie	THC:CBD im Verhältnis 1:1	Placebo	Vor- und Weiterbehandlung,
GWMS0001	<p>GW-1000-02: THC (27 mg/ml)/CBD (25 mg/ml) mit Pfefferminzöl 0,05 % (v/v) als Spray zur Anwendung in der Mundhöhle.</p> <p>GW-1001-02: THC (27mg/ml)/CBD (25 mg/ml) 0,05 % (v/v) als Spray zur Anwendung in der Mundhöhle</p> <p>Jeder Sprühstoß gibt 0,1 ml ab; maximal 8 Sprühstöße über 3 Stunden, maximal 48 Sprühstöße über 24 Stunden</p>	<p>GW-4001-01 (Placebo) Pfefferminzöl 0,05 %, gelbes Chinolin 0,005 % (w/v), Sunset Yellow 0,0025 % (w/v)</p> <p>Jeder Sprühstoß gibt 0,1 ml ab; maximal 8 Sprühstöße über 3 Stunden, maximal 48 Sprühstöße über 24 Stunden</p>	Die medikamentöse Vorbehandlung der Spastik wurde im Studienzeitraum fortgesetzt (Add-on Gabe der Studienmedikation).
GWMS0106	<p>THC (27 mg/ml)/CBD (25 mg/ml) mit Pfefferminzölgeschmack 0.05 % (v/v) als Spray zur Anwendung in der Mundhöhle.</p> <p>Jeder Sprühstoß gibt 0,1 ml ab; maximal 8 Sprühstöße über 3 Stunden, maximal 48 Sprühstöße über 24 Stunden</p>	<p>Placebo Pfefferminzöl, 0.05 % (v/v), gelbes Chinolin 0.005 % (w/v) als Spray zur Anwendung in der Mundhöhle.</p> <p>Jeder Sprühstoß gibt 0,1 ml ab; maximal 8 Sprühstöße über 3 Stunden, maximal 48 Sprühstöße über 24 Stunden</p>	Die medikamentöse Vorbehandlung der Spastik wurde im Studienzeitraum fortgesetzt (Add-on Gabe der Studienmedikation).
GWCL0403	<p>THC (27 mg/ml)/CBD (25 mg/ml) mit Pfefferminzölgeschmack 0.05 % (v/v) als Spray zur Anwendung in der Mundhöhle.</p> <p>Jeder Sprühstoß gibt 0,1 ml ab; maximal 8 Sprühstöße über 3 Stunden, maximal 24 Sprühstöße über 24 Stunden</p>	<p>Placebo Pfefferminzöl, 0.05 % (v/v), gelbes Chinolin 0.005 % (w/v) als Spray zur Anwendung in der Mundhöhle.</p> <p>Jeder Sprühstoß gibt 0,1 ml ab; maximal 8 Sprühstöße über 3 Stunden, maximal 24 Sprühstöße über 24 Stunden</p>	Die medikamentöse Vorbehandlung der Spastik wurde im Studienzeitraum fortgesetzt (Add-on Gabe der Studienmedikation).
GWSP0604	<p>Run-in Phase: THC (27 mg/ml)/CBD (25 mg/ml) mit Pfefferminzölgeschmack 0.05 % (v/v) als Spray zur Anwendung in der Mundhöhle.</p> <p>Jeder Sprühstoß gibt 0,1 ml ab; 12 Sprühstöße je 24 Stunden</p> <p>Randomisierte Weiterbehandlung: THC (27 mg/ml)/CBD (25</p>	<p>Run-in Phase: THC (27 mg/ml)/CBD (25 mg/ml) mit Pfefferminzölgeschmack 0.05 % (v/v) als Spray zur Anwendung in der Mundhöhle.</p> <p>Jeder Sprühstoß gibt 0,1 ml ab; 12 Sprühstöße je 24 Stunden</p> <p>Randomisierte Weiterbehandlung: Placebo Pfefferminzöl,</p>	Die medikamentöse Vorbehandlung der Spastik wurde im Studienzeitraum fortgesetzt (Add-on Gabe der Studienmedikation).

Studie	THC:CBD im Verhältnis 1:1	Placebo	Vor- und Weiterbehandlung,
	mg/ml) mit Pfefferminzölgeschmack 0.05 % (v/v) als Spray zur Anwendung in der Mundhöhle. Jeder Sprühstoss gibt 0,1 ml ab; maximal 12 Sprühstöße je 24 Stunden	0.05 % (v/v), gelbes Chinolin 0.005 % (w/v) als Spray zur Anwendung in der Mundhöhle. Jeder Sprühstoss gibt 0,1 ml ab; maximal 12 Sprühstöße je 24 Stunden	
ARA2009	THC (27 mg/ml)/CBD (25 mg/ml) mit Pfefferminzölgeschmack 0.05 % (v/v) als Spray zur Anwendung in der Mundhöhle. Jeder Sprühstoß gibt 0,1 ml ab; maximal 8 Sprühstöße über 3 Stunden, maximal 48 Sprühstöße über 24 Stunden	Placebo Pfefferminzöl, 0.05 % (v/v), gelbes Chinolin 0.005 % (w/v) als Spray zur Anwendung in der Mundhöhle. Jeder Sprühstoß gibt 0,1 ml ab; maximal 8 Sprühstöße über 3 Stunden, maximal 48 Sprühstöße über 24 Stunden	Die medikamentöse Vorbehandlung der Spastik wurde im Studienzeitraum fortgesetzt (Add-on Gabe der Studienmedikation).

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

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

Studie Gruppe	N	Alter [Jahre] (SD)	Geschlecht w /m (%)	Dauer der MS- Erkrankung [Jahre] MW (±SD)	Dauer der Spastik- erkrankung [Jahre] MW (±SD)	Vorange- gangener Cannabis Konsum n (%)	Medikamentöse Begleitbehandlung der Spastik [Anzahl der Patienten] (%)				
							antispasti- sche Medikati- on insgesamt	Baclofen	Tiza- nidin	Benzo- diazepine	Dantro- len
GWMS0001											
GW-1000-02/ GW-1001-02	n= 80	51,0 (9,4)	47 (58,8)/ 33 (41,2)	NR	NR	43 (53,8)	78 (97,5)	NR	NR	11 (13,8)	1 (1,3)
GW-4001-01 (Placebo)	n= 80	50,4 (9,3)	52 (65,0)/ 28 (35,0)	NR	NR	53 (66,3)	75 (93,8)	NR	NR	1 (1,3)	0
GWMS0106											
Sativex®	n= 124	49,7 (10,2)	80 (64,5)/ 44 (35,5)	13,61 (8,605)	NR	52 (41,9)	93 (75)	56 (45)	27 (22)	25 (21)	3 (3)
Placebo	n= 65	47,8 (9,5)	34 (52,3)/ 31 (47,7)	12,19 (7,687)	NR	27 (41,5)	42 (65)	24 (37)	13 (20)	13 (21)	4 (6)
GWCL0403											
Sativex®	n= 167	48,0 (10,06)	106 (63)/ 61 (37)	14,4 (8,28)	7,5 (5,14)	34 (20)	160 (96)	132 (79)	69 (41)	44 (26)	14 (8)
Placebo	n= 170	47,1 (9,15)	101 (59)/ 69 (41)	16,0 (8,48)	8,0 (5,51)	47 (28)	166 (98)	138 (81)	76 (45)	51 (30)	10 (6)
GWSP0604											
Phase A											
Sativex®	n = 572	49,1 (9,85)	129 (39)/ 202 (61)	12,3 (7,49)	7,4 (5,54)	43 (13)	488 (85)	401 (70)	ND	110 (19)	4 (1)

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

Studie Gruppe	N	Alter [Jahre] (SD)	Geschlecht w /m (%)	Dauer der MS-Erkrankung [Jahre] MW (±SD)	Dauer der Spastikerkrankung [Jahre] MW (±SD)	Vorangegangener Cannabis Konsum n (%)	Medikamentöse Begleitbehandlung der Spastik [Anzahl der Patienten] (%)				
							antispastische Medikation insgesamt	Baclofen	Tizanidin	Benzodiazepine	Dantrolen
Phase B											
Sativex®	n= 124	49,1 (9,09)	72 (58)/ 52 (42)	13.3 (8.29)	8.6 (6.89)	22 (18)	111 (90)	87 (70)	ND	23 (18)	1 (1)
Placebo	n= 117	48,1 (9,59)	73 (62)/ 44 (38)	11.8 (7.38)	6.7 (5.40)	22 (19)	103 (88)	90 (77)		30 (25)	0
ARA2009											
Sativex®	n=17	49,8 (6,64)	6 (35)/ 11 (65)	20,76 (8,42)	ND	ND	10 (ND)	ND	ND	ND	ND
Placebo	n=17	49,8 (6,64)	6 (35)/ 11 (65)	20,76 (8,42)	ND	ND	10 (ND)	ND	ND	ND	ND

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

Alle RCT waren multizentrisch (außer ARA2009), randomisiert, doppelblind und Placebo-kontrolliert und untersuchten die Wirkstoffkombination THC:CBD im Verhältnis 1:1 in der Zusatzbehandlung von Patienten mit MS-bedingter Spastik, wobei in der Studie GWMS0001 nur 24 % der Patienten die Spastik als primäres Symptom angaben (37 von 160 Patienten). Allerdings waren insgesamt 87,5 % der Patienten von spastischen Symptomen betroffen und erhielten dementsprechende Medikation (140 von 160 Patienten).

Die Studien GWMS0001, GWMS0106 und GWCL0403 untersuchten die Zusatztherapie von Patienten, die in der Vortherapie hinsichtlich der Wirkstoffkombination THC:CBD 1:1 naive waren. In den Studien GWMS0001 und GWMS0106 erfolgte eine aktive Behandlungsphase von 6 Wochen und bei GWCL0403 eine aktive Behandlungsphase von 14 Wochen zur Erfassung der Behandlungseffekte.

Die Studie GWSP0604 war die konfirmatorische Wirksamkeitsstudie unter Respondern, d.h., dass nur Patienten untersucht wurden, die in einem vorgeschalteten einfach-verblindeten 4-wöchigen Therapieversuch auch tatsächlich mit einer klinisch relevanten Verbesserung der Symptomatik auf die Behandlung mit der Wirkstoffkombination THC:CBD 1:1 reagiert hatten. Abbildung 3 gibt eine Übersicht über die Art und Dauer der Behandlungsphasen und die Erhebungszeitpunkte dieser Studie.

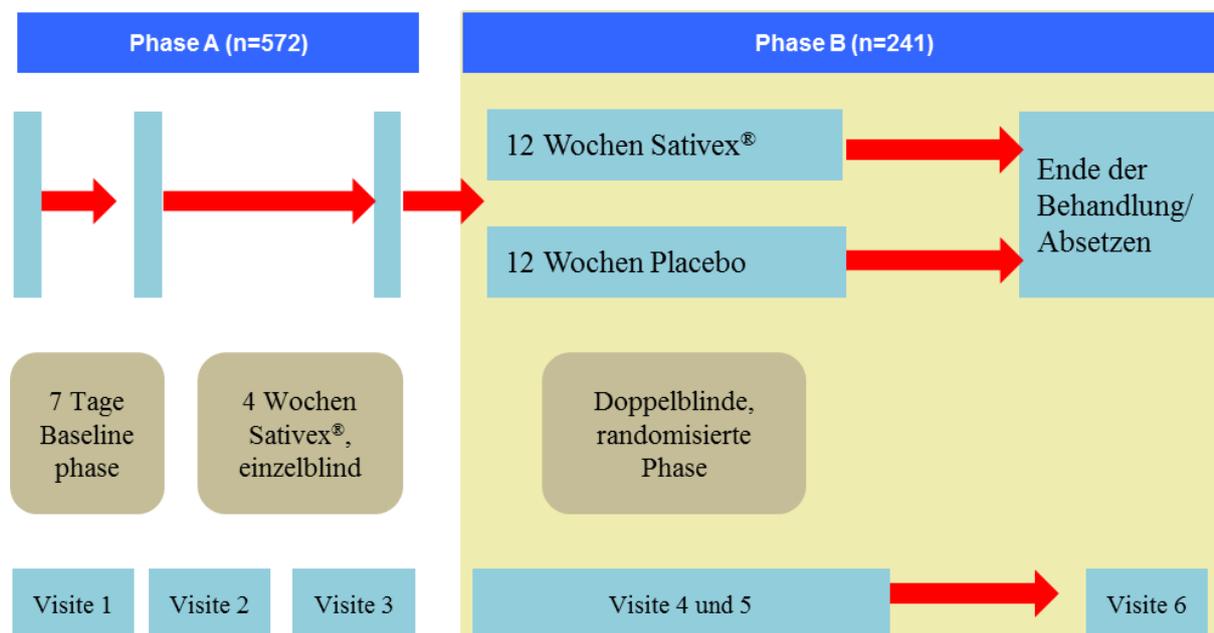


Abbildung 3: Art und Dauer der Behandlungsphasen und der Erhebungszeitpunkte der konfirmatorischen Wirksamkeitsstudie GWSP0604.

Die Studie ARA2009 war eine explorative Studie mit 34 Patienten, die eine Subpopulation von Patienten mit Spastik aus einer größeren MS-Studienpopulation hinsichtlich der Wirkung einer 3-wöchigen Behandlung mit einem Cannabinoid-Extrakt (THC/CBD etwa 1:1) auf verschiedene psychopathologische Dimensionen, die Verträglichkeit sowie die Lebensqualität untersuchte.

Die Studienpopulationen der einzelnen Studien zeigten ähnliche Charakteristika hinsichtlich des Durchschnittsalters von um die 49 Jahre, der Geschlechterverteilung mit einem höheren Frauenanteil und der Dauer der MS Erkrankung von im Durchschnitt 11 Jahren oder länger, soweit bekannt. In der Studie ARA2009 war die Vorerkrankungsdauer der MS mit durchschnittlich über 20 Jahren im Vergleich aller Studien am höchsten.

In allen RCT waren die Patienten mehrheitlich hinsichtlich der Spastik medikamentös vorbehandelt (Tabelle 4-9):

Tabelle 4-9: Medikamentöse Vorbehandlung und Cannabis-Konsum

Studie	Patienten, die anti-spastische Medikation erhielten [%]		Patienten, die vorher Cannabis konsumierten [%]	
	THC:CBD 1:1	Placebo	THC:CBD 1:1	Placebo
GWMS0001*	98	94	54	67
GWMS0106	75	65	42	42
GWCL0403	96	100	20	28
GWSP0604	90	88	18	19
ARA2009	59	59	NR	NR

* Gesamtpool der an der Studie teilnehmenden Patienten

In den Studien GWMS0001, GWCL0403 und GWSP0604 waren nahezu alle Patienten medikamentös vortherapiert. In der Studie ARA2009 erhielten 10 der 17 Patienten eine anti-spastische Vortherapie (Baclofen, Dantrolen, Tizanidin oder Botulinumtoxin).

Für die Studie GWMS0001 konnte die Vortherapie der 37 Patienten, mit Spastik als primärem Symptom, von 160 Patienten nicht ermittelt werden. In allen RCT war die häufigste Begleittherapie der Spastik Baclofen gefolgt von Tizanidin und Benzodiazepinen. Die Behandlung mit Dantrolen war vergleichsweise selten.

Die Pharmakokinetik des psychoaktiven Cannabinoids THC unterscheidet sich grundlegend dahingehend, ob es als Spray verabreicht wird, wie es bei Sativex® der Fall ist, oder ob es inhaliert wird. Bei Sativex® wird eine maximale Konzentration von 3 ng/ml innerhalb einer

T_{\max} von 60 bis 90 Minuten erreicht. Im Gegensatz dazu werden durch das Rauchen von Cannabis maximale Konzentrationen von 100 ng/ml innerhalb weniger Minuten erzielt(39).

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 4-10: 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			
GWMS0001	ja	ja	ja	ja	nein	nein	niedrig
GWMS0106	ja (2:1)	ja	ja	ja	nein	nein	niedrig
GWCL0403	ja (1:1)	ja	ja	ja	nein	nein	niedrig
GWSP0604							
Phase B	ja (1:1)	ja	ja	ja	nein	nein	niedrig
ARA2009	ja (1:1, cross-over)	unklar	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Alle Studien wurden in den Phasen der Gruppenvergleiche doppelblind durchgeführt. Das gilt auch für die Responderstudie GWSP0604, die lediglich in der Run-in-Phase, in der alle Patienten Verum (THC:CBD 1:1) erhielten, einfach verblindet war. In allen fünf Studien wurde die Randomisierung von einer unabhängigen Personen computer-basiert zentral durchgeführt. Je nach Studie wurde die Zuteilung der Patienten zu den Studiengruppen entweder über ein IVRS oder konsekutiv durch das verblindete Personal im Zentrum vorgenommen. Die Verdeckung der Gruppenzuteilung in der Studie ARA2009 war unklar. Die statistischen Analysen wurden wie geplant durchgeführt und alle Ergebnisse wurden vollständig im Clinical Study Report dargestellt. Die Publikationen der Ergebnisse aus den Studien berichten sowohl primäre und sekundäre Endpunkte. Insofern ergeben sich insgesamt keine Hinweise auf ein hohes Verzerrungspotential auf Studienebene für alle eingeschlossenen RCT.

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 4-11: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Schweregrad der Spastik (NRS)	Anzahl Patienten mit klinisch relevanter Verbesserung der Schwere der Spastik	Anzahl der Spastiken pro Tag	Schlafqualität	Gesundheitszustand (EQ-5D)
GWMS0001	Ja	Ja	Nein	Ja	Nein
GWMS0106	Ja	Ja	Ja	Nein	Nein
GWCL0403	Ja	Ja	Nein	Ja	Ja
GWSP0604	Ja	Ja	Ja	Ja	Ja
ARA2009	Nein	Nein	Nein	Nein	Nein
Studie	Lebensqualität (MSQoL, SF-36 oder VAS QoL)	Aktivitäten des täglichen Lebens (Barthel Index)	Gehfähigkeit (10 m-Gehstreckentest)	Klinische Globalbeurteilung (Patient)	Klinische Globalbeurteilung (betreuende Person)
GWMS0001	Nein	Ja	Ja	Nein	Nein
GWMS0106	Nein	Nein	Nein	Ja	Nein
GWCL0403	Ja	Ja	Ja	Nein	Ja
GWSP0604	Ja	Ja	Ja	Ja	Ja
ARA2009	Ja	Nein	Nein	Nein	Nein
Studie	Klinische Globalbeurteilung (Arzt)	Nebenwirkungen			
GWMS0001	Nein	Ja			
GWMS0106	Nein	Ja			
GWCL0403	Nein	Ja			
GWSP0604	Ja	Ja			
ARA2009	Nein	Ja			

4.3.1.3.1 <Endpunkt xxx> – 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.

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.

4.3.1.3.1.1 Absolute Veränderung des Grads der Spastizität auf der NRS – RCT

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

Tabelle 4-12: Operationalisierung von Absolute Veränderung des Grads der Spastizität auf der NRS

Studie	Operationalisierung
GWMS0001	Der Schweregrad der Spastik wurde von den Patienten im Tagebuch über die gesamte Studiendauer (1-2 Wochen Screening plus 6 Wochen Behandlungsphase) oder bis zum vorzeitigen Studienende dokumentiert. Im Tagebuch wurde die durchschnittliche Schwere der Spastik über die vergangenen 24 Stunden auf einer VAS (0-100) dokumentiert. Der absolute Rückgang der Schwere der Spastik wurde als Änderung des 7-Tage Durchschnitts der VAS im Vergleich zum Baseline-Wert analysiert. Die Werte der VAS (0-100 mm) wurden im Rahmen dieser Nutzenbewertung linear auf die 11-stufige NRS (0-10) umgerechnet.
GWMS0106	Der Schweregrad der Spastik wurde von den Patienten im Tagebuch dokumentiert über die gesamte Studiendauer von 8 Wochen (2 Wochen Screeningphase plus 6 Wochen Behandlungsphase) oder bis zum vorzeitigen Studienende. Im Tagebuch wurde die durchschnittliche Schwere der Spastik über die vergangenen 24 Stunden auf einer 11-stufigen NRS (0-10) dokumentiert (0= 'no spasticity', 10= 'worst ever spasticity'). Der absolute Rückgang der Schwere der Spastik wurde analysiert als Änderung des 7 Tagedurchschnitts der NRS zum Studienende im Vergleich zu Baseline.
GWCL0403	Der Schweregrad der Spastik wurde von den Patienten im Tagebuch dokumentiert über die gesamte Studiendauer von 15 Wochen (1 Woche Screeningphase plus 14 Wochen Behandlungsphase mit Visiten in Woche -1, 0, 2, 6, 10 und 14) oder bis zum vorzeitigen Studienende. Im Tagebuch wurde die durchschnittliche Schwere der Spastik über die vergangenen 24 Stunden auf einer 11-stufigen NRS (0-10) dokumentiert (0= 'no spasticity', 10= 'worst ever spasticity'). Der absolute Rückgang der Schwere der Spastik wurde analysiert als Änderung des 14 Tagedurchschnitts der NRS zum Studienende im Vergleich zu Baseline.
GWSP0604	Der Schweregrad der Spastik wurde von den Patienten im Tagebuch dokumentiert über die gesamte Studiendauer (1 Woche Screeningphase plus 4 Wochen Studienmedikation Run-in Phase plus 12 Wochen randomisierte Fortsetzung versus Absetzen der Studienmedikation bei Ansprechen auf die Studienmedikation in der Run-in Phase oder bis zum vorzeitigen Studienende). Im Tagebuch wurde die durchschnittliche Schwere der Spastik über die vergangenen 24 Stunden auf einer 11-stufigen NRS (0-10) dokumentiert (0= 'no spasticity', 10= 'worst ever spasticity'). Der absolute Rückgang der Schwere der Spastik wurde analysiert als Änderung des 14 Tagedurchschnitts der NRS zum Studienende im Vergleich zu Baseline.

Die Spastizität wurde in einer Studie (31) mittels der 100 mm Visuellen Analogskala (VAS) und in den anderen Studien mit Hilfe der 0-10 Numerischen Ratingskala (NRS) bestimmt. Daten der 0-100 mm VAS wurden im Rahmen dieser Nutzenbewertung durch einfache lineare Transformation in die 0-10 NRS konvertiert, in dem jede aufgezeichnete individuelle Beobachtung durch 10 geteilt wurde (3), um auch diese Ergebnisse in der Meta-Analyse berücksichtigen zu können.

Es wird davon ausgegangen, dass die Änderung des NRS, die im Vergleich zur Baseline im 7 oder 14 Tagesdurchschnitt erfolgte, keinen Einfluss auf das Ergebnis des Endpunktes Veränderung des Grads der Spastizität hat.

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 4-13: Bewertung des Verzerrungspotenzials für Absolute Veränderung des Grads der Spastizität auf der NRS 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
GWMS0001	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWMS0106	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWCL0403	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWSP0604	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Aufgrund des niedrigen Verzerrungspotentials auf Studienebene aller relevanten RCT, die den Endpunkt Schweregrad Spastizität auf der 11-stufigen NRS bzw. der VAS untersucht haben, kann von einem niedrigen Verzerrungspotential auch auf Ebene des Endpunkts selbst ausgegangen werden. In den vier Studien waren die Endpunkterheber (Patienten) verblindet. Der Endpunkt wurde wie geplant für die jeweilige ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Die Studie GWMS0001 wurde entsprechend der im Dossier verwendete Subgruppe von 37 Patienten (von insgesamt 160), mit Spastizität als primäres MS-Symptom, als ITT-Population analysiert. Hinweise auf eine ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt wie geplant statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in allen vier Studien das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

Stellen Sie die Ergebnisse für den Endpunkt Veränderung des Grads der Spastizität für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-14: Ergebnisse für Absolute Veränderung des Grads der Spastizität auf der NRS aus RCT mit dem zu bewertenden Arzneimittel

Studie	Veränderung des Grads der Spastizität (NRS Score)								
	Baseline			Studienende			Veränderung		
	N	Mean (SD)	Median Min-Max	N	Mean (SD)	Median Min-Max	Mean (SD)	Median Min-Max	Gruppen- differenz (*) (95%-CI) p-Wert
GWMS 0001 ¹⁾									
GW-1000-02	20	6,00 (1,92)	6,55 (0,8-9,0)	19	4,49 (2,14)	4,80 (1,0-7,7)	-1,43 (2,25)	-0,14 (-5,5- -2,7)	-1,84 (-3,181; - 0,501)
Placebo	19	6,99 (1,58)	7,70 (4,3-9,6)	18	6,97 (2,02)	7,68 (1,1-9,1)	-0,15 (1,79)	-0,15 (-4,0- 2,8)	p=0,009
GWMS 0106 ¹⁾									
Sativex®	122	5,49 (1,91)	5,43 (0,3- 9,2)	122	4,31 (2,25)	4,14 (0,1-10)	-1,18 (1,82)	-1,0 (-6,7-2,7)	-0,517 (-1,029; - 0,004)
Placebo	64	5,39 (1,91)	5,57 (1,1- 9,9)	65	4,75 (2,25)	4,86 (0,2-10)	-0,63 (1,61)	-0,23 (-7,7-2,6)	p=0,048
GWCL 0403 ²⁾									
Sativex®	166	6,77 (1,33)	6,83 (3,0-10)	166	5,61 (2,03)	5,7 (0,5-10)	-1,16 (1,78)	-0,8 (-8,0-3,8)	-0,17 (-0,53; 0,19)
Placebo	169	6,48 (1,31)	6,5 (3,5-9,5)	166	5,58 (1,96)	5,64 (0,9-9,5)	-0,91 (1,70)	-0,67 (-6,1-4,1)	p=0,3592
GWSP 0604 ³⁾									
Sativex®	124	3,87 (1,49)	4,0 (0-6,6)	124	3,68 (1,83)	3,71 (0-9,0)	-0,19 (1,35)	-0,21 (-5,4-4,4)	-0,84 (-1,29; -0,4)
Placebo	117	3,92 (1,55)	4,0 (0-7,1)	117	4,56 (2,31)	4,29 (0,2-9,0)	0,64 (2,14)	0,29 (-5,3-7,1)	p=0,0002
(*) Adjustiertes ANCOVA Modell									
1. Verbesserung 6 Wochen nach Behandlungsbeginn									
2. Verbesserung 14 Wochen nach Behandlungsbeginn									
3. Verbesserung 12 Wochen nach Fortführung versus Absetzen nach Ansprechen (mindestens 20 % Verbesserung der NRS) auf 4 Wochen Run-in Behandlung mit Sativex®									

Im direkten Vergleich zur Therapiefortsetzung ohne die Wirkstoffkombination THC:CBD 1:1 zeigt die Add-on Therapie mit der Wirkstoffkombination THC:CBD 1:1 in allen vier RCT eine deutlichere Verbesserung der Spastik auf der 11-stufigen NRS bzw. der 0-100 mm VAS. In drei der vier Studien (GWMS0001, GWMS0106 und GWSP0604) war die Mittelwertdifferenz im Vergleich zur Therapiefortsetzung ohne die Wirkstoffkombination THC:CBD 1:1 statistisch signifikant zu Gunsten der Zusatztherapie mit der Wirkstoffkombination THC:CBD 1:1. Die Studie GWCL0403 zeigte numerische Vorteile von THC:CBD 1:1, die statistisch jedoch nicht signifikant sind. Ein Grund dafür könnte in der hohen Abbruchrate innerhalb der beiden Behandlungsarme (17 Sativex[®], 15 Placebo) liegen. Insgesamt 14 Patienten brachen die Studie aufgrund von Nebenwirkungen ab (9 Sativex[®], 5 Placebo). Weitere Gründe lagen gleich verteilt in beiden Behandlungsarmen in der Rücknahme der Einverständniserklärung, dem Fehlen der Wirksamkeit und dem Lost-to-follow up.

Wie in nachfolgender Tabelle dargestellt (Tabelle 4-15), zeigt die Analyse der Per Protocol (PP) Population eine statistisch signifikante Verbesserung des Grads der Spastik, gemessen auf der NRS.

Tabelle 4-15: PP-Analyse der Studie GWCL0403 für Absolute Veränderung des Grads der Spastizität auf der NRS

	N	Adjustierte Differenz zwischen Baseline und Studienende	Gruppendifferenz [95%-CI]	p-Wert
Sativex [®]	122			
		-0,84	-0,46 [-0,88; -0,03]	0,0348
Placebo	143			

In der Studie, die dem Zulassungsstatus von Sativex[®] am besten entspricht, also den Patienten, die auf eine Therapie mit Sativex[®] in klinisch relevanter Form ansprechen, (Responder Studie GWSP0604) wurde statistische Signifikanz erzielt.

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.

Es wurde eine Meta-Analyse für drei Studien durchgeführt (GWMS0001, GWMS0106 und GWCL0403), da diese drei Studien hinsichtlich dieses Endpunkts vergleichbar waren. Die Meta-Analyse wurde mit dem Programm RevMan 5.1 durchgeführt.

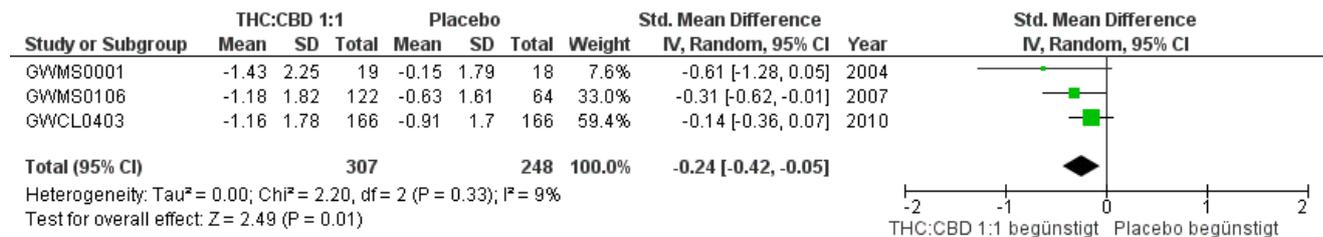


Abbildung 4: Meta-Analyse für den Endpunkt Veränderung des Grads der Spastizität

In der Meta-Analyse betrug die Mittelwertdifferenz des NRS Scores der Spastikbeurteilung -0,24 (95 %-CI: -0,42; -0,05) zugunsten der Wirkstoffkombination THC:CBD 1:1 und zeigte statistische Signifikanz auf dem 5 % Niveau ($p = 0,01$). Somit zeigt sich in der Meta-Analyse dieser drei RCT auch in der Mischpopulation von Respondern und Non-Respondern eine signifikante Abnahme der Spastizität auf der NRS (Abbildung 4). Beschränkt man sich auf die Zielpopulation lt. Fachinformation, d.h. scheiden die Non-Responder nach 4 Wochen Therapiever such aus, ist ein deutlich höherer Wert zu erwarten. Das zeigen die Ergebnisse der Studie GWSP0604.

Die Studie GWSP0604 wurde nicht in eine erweiterte Meta-Analyse integriert, da diese RCT eine andere Patientenpopulation untersuchte. Bei diesen Patienten wurde die Wirksamkeit von Sativex[®] bei Patienten untersucht, die auf einen Therapiever such mit Sativex[®] bereits positiv angesprochen hatten.

Da in den drei Studien, die in der Meta-Analyse untersucht wurden, die Ergebnisse nicht nach Respondern und Non-Respondern getrennt vorliegen, kann keine Sensitivitätsanalyse zusammen mit den Ergebnissen aus der Studie GWSP0604 durchgeführt werden.

4.3.1.3.1.2 Anzahl Patienten mit klinisch relevantem Rückgang der Schwere der Spastizität (Responder) auf der NRS – 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.

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

Tabelle 4-16: Operationalisierung von Anzahl Patienten mit klinisch relevantem Rückgang der Schwere der Spastizität (Responder)

Studie	Operationalisierung
GWMS0001	Der Schweregrad der Spastik wurde von den Patienten im Tagebuch über die gesamte Studiendauer (1-2 Wochen Screening plus 6 Wochen Behandlungsphase) oder bis zum vorzeitigen Studienende dokumentiert. Im Tagebuch wurde die durchschnittliche Schwere der Spastik über die vergangenen 24 Stunden auf einer VAS (0-100) dokumentiert. Der absolute Rückgang der Schwere der Spastik wurde als Änderung des 7-Tage Durchschnitts der VAS im Vergleich zum Baseline-Wert analysiert. Die Werte der VAS (0-100 mm) wurden im Rahmen dieser Nutzenbewertung linear auf die 11-stufige NRS (0-10) umgerechnet.
GWMS0106	Der Schweregrad der Spastik wurde von den Patienten im Tagebuch dokumentiert über die gesamte Studiendauer von 8 Wochen (2 Wochen Screeningphase plus 6 Wochen Behandlungsphase) oder bis zum vorzeitigen Studienende. Im Tagebuch wurde die durchschnittliche Schwere der Spastik über die vergangenen 24 Stunden auf einer 11-stufigen NRS (0-10) dokumentiert (0= 'no spasticity', 10= 'worst ever spasticity'). Der Rückgang der Schwere der Spastik wurde patientenbezogen analysiert als Änderung des 7 Tage NRS zum Studienende um mindestens 30 % oder mindestens 50 % im Vergleich zu Baseline.
GWCL0403	Der Schweregrad der Spastik wurde von den Patienten im Tagebuch dokumentiert über die gesamte Studiendauer von 15 Wochen (1 Woche Screeningphase plus 14 Wochen Behandlungsphase mit Visiten in Woche -1, 0, 2, 6, 10 und 14) oder bis zum vorzeitigen Studienende. Im Tagebuch wurde die durchschnittliche Schwere der Spastik über die vergangenen 24 Stunden auf einer 11-stufigen NRS (0-10) dokumentiert (0= 'no spasticity', 10= 'worst ever spasticity'). Der Rückgang der Schwere der Spastik wurde patientenbezogen analysiert als Änderung des 14 Tage NRS nach Studienenden um mindestens 30 % oder mindestens 50 % im Vergleich zu Baseline.
GWSP0604	Der Schweregrad der Spastik wurde von den Patienten im Tagebuch dokumentiert über die gesamte Studiendauer (1 Woche Screeningphase plus 4 Wochen Studienmedikation Run-in Phase plus 12 Wochen randomisierte Fortsetzung versus Absetzen der Studienmedikation bei Ansprechen auf die Studienmedikation in der Run-in Phase oder bis zum vorzeitigen Studienende). Im Tagebuch wurde die durchschnittliche Schwere der Spastik über die vergangenen 24 Stunden auf einer 11-stufigen NRS (0-10) dokumentiert (0= 'no spasticity', 10= 'worst ever spasticity'). Der Rückgang der Schwere der Spastik wurde patientenbezogen analysiert als Änderung des 14 Tage NRS nach Studienende um mindestens 30 % oder mindestens 50 % im Vergleich zu Baseline.

In allen vier Studien wurde der Schweregrad der Spastik von den Patienten dokumentiert. In der Studie GWMS0001 wurde die Spastik auf der VAS dokumentiert und im Rahmen dieser Nutzenbewertung auf die NRS umgerechnet.

Es wird davon ausgegangen, dass die Änderung des NRS, die im Vergleich zur Baseline im 7 oder 14 Tagesdurchschnitt erfolgte, keinen Einfluss auf das Ergebnis des Endpunktes klinisch relevanter Rückgang der Schwere der Spastizität hat.

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 4-17: Bewertung des Verzerrungspotenzials für Anzahl Patienten mit klinisch relevantem Rückgang der Schwere der Spastizität (Responder) 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
GWMS0001	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWMS0106	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWCL0403	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWSP0604	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Aufgrund des niedrigen Verzerrungspotentials auf Studienebene aller relevanten RCT, die den Endpunkt „Anzahl der Patienten mit klinisch relevantem Rückgang der Schwere der Spastizität“ untersucht haben, kann von einem niedrigen Verzerrungspotential auch auf Ebene des Endpunkts selbst ausgegangen werden. In den vier Studien waren die Endpunkterheber (Patienten) verblindet. Der Endpunkt wurde, wie geplant, für die jeweilige ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf eine ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt wie geplant statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in allen vier Studien das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

Stellen Sie die Ergebnisse für den Endpunkt klinisch relevanter Rückgang der Schwere der Spastizität (Responder) für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-18: Ergebnisse für Anzahl Patienten mit klinisch relevantem Rückgang der Schwere der Spastizität (Responder, 30% Verminderung) gemessen mittels NRS aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anteil Patienten mit ≥ 30 % Verbesserung der Schwere der Spastik				Anteil Patienten mit ≥ 50 % Verbesserung der Schwere der Spastik		
	N	n (%)	Differenz [%] (95%-CI) p-Wert	Odds ratio (95%-CI) p-Wert	n (%)	Differenz [%] (95%-CI) p-Wert	Odds ratio (95%-CI) p-Wert
GWMS 0001 ¹⁾							
GW- 1000-02	19	12 (63,2)	46,49 (18,8;74,18)	8,57 (1,82; 40,42)	5 (26,13)	15,2 (-9,35; 39,76)	3,04 (0,51; 18,11)
Placebo	18	3 (16,7)	ND	p=0,007	2 (11,1)	ND	p=0,22
GWMS 0106 ¹⁾							
Sativex®	120	48 (40,0)	18,13 (4,73; 31,52)	2,38 (1,19; 4,78)	21 (17,5)	8,13 (-1,73; 17,98)	2,05 (0,78; 5,37)
Placebo	64	14 (21,9)	p= 0,014	p=0,01	6 (9,4)	p= 0,189	p=0,14
GWCL 0403 ²⁾							
Sativex®	166	51 (31,0)	5,9 (-3,7; 15,4)	1,34 (0,83; 2,17)	21 (13,0)	2,0 (-4,9; 8,9)	1,21 (0,62; 2,37)
Placebo	169	42 (25,0)	ND	p=0,23	18 (11,0)	ND	p=0,569
GWSP 0604 ³⁾							
Sativex®	124	92 (74,0)	22,9 (11,0; 34,88)	2,73 (1,59; 4,69)	56 (45,0)	11,8 (-0,4; 24,1)	1,65 (0,98; 2,78)
Placebo	117	60 (51,0)	ND	p=0,0003	39 (33,0)	ND	p=0,061
(1) Verbesserung 6 Wochen nach Behandlungsbeginn (2) Verbesserung 14 Wochen nach Behandlungsbeginn (3) Verbesserung 12 Wochen nach Fortführung versus Absetzen nach Ansprechen (mindestens 20% Verbesserung der NRS) auf 4 Wochen Run-in Behandlung mit Sativex®							

Im direkten Vergleich zur Therapiefortsetzung ohne die Wirkstoffkombination THC:CBD 1:1 zeigt die Add-on Therapie unter der Wirkstoffkombination THC:CBD 1:1 in allen vier RCT eine höhere Zahl von Patienten mit einer klinisch relevanten Verbesserung der Spastik (30 % oder mehr Verbesserung auf der 11-stufigen NRS). Bei naiven Patienten lag die Ansprechrate unter THC:CBD 1:1 in der Studie GWMS0001 bei 63,2 % (Placebo 16,7 %), resultierend in

einer Odds ratio von 8,57 (95%-CI: 1,82; 40,42), in der Studie GWMS0106 bei 40 % (Placebo: 21,9 %), resultierend in einer Odds ratio von 2,38 (95 %-CI: 1,19; 4,78) und in der Studie GWCL0403 unter THC:CBD 1:1 bei 31 % (Placebo: 25 %), resultierend in einer Odds ratio von 1,34 (95 %-CI: 0,83;2,17).

Eine Erklärung für den statistisch nicht signifikanten Unterschied für die Studie GWCL0403 für diesen Endpunkt könnte in der hohen Abbruchrate innerhalb der beiden Behandlungsarme (17 Sativex[®], 15 Placebo) liegen. Insgesamt 14 Patienten brachen die Studie aufgrund von Nebenwirkungen ab (9 Sativex[®], 5 Placebo). Weitere Gründe lagen gleichverteilt in beiden Behandlungsarmen in der Rücknahme der Einverständniserklärung, dem Fehlen der Wirksamkeit und dem Lost-to-follow up.

Wie in nachfolgender Tabelle dargestellt (Tabelle 4-19), zeigt die Analyse der Per Protocol (PP) Population eine statistisch signifikant höhere Zahl an Patienten mit klinisch relevanter Verbesserung der Schwere der Spastik, gemessen auf der NRS. Die Odds ratio beträgt in der Betrachtung der PP-Population der Responder mit mindestens 30 % Verbesserung 1,74 (95 %-CI: 1,02; 2,96).

Tabelle 4-19: PP-Analyse der Studie GWCL0403 für Anzahl Patienten mit klinisch relevantem Rückgang der Schwere der Spastizität gemessen mittels NRS

	N	Anteil Patienten mit ≥ 30 % Verbesserung der Schwere der Spastik			Anteil Patienten mit ≥ 50 % Verbesserung der Schwere der Spastik		
		n (%)	Differenz [%] (95%-CI)	Odds ratio (95%-CI) p-Wert	n (%)	Differenz [%] (95%-CI)	Odds ratio (95%-CI) p-Wert
Sativex [®]	122	44 (36)	11,6 (5,0; 22,6)	1,74 (1,02; 2,96) p=0,0402	20 (16)	5,9 (-2,4; 12,2)	1,67 (0,82; 3,43) p=0,1579
Placebo	143	35 (24)			15 (10)		

In der Responderstudie GWSP0604 lag die Ansprechrate von ≥ 30 % unter Patienten, die positiv auf eine 4-wöchige Vortherapie mit Sativex[®] angesprochen hatten, nach Fortsetzung der Therapie bei 74 % (Placebo 51,0 %) resultierend in einer statistisch signifikanten Odds ratio von 2,73 (95 %-CI: 1,59; 4,69). Basierend auf dem Kriterium für Ansprechen von mindestens 50 % Verbesserung auf der NRS lag der Anteil der Patienten mit Ansprechen unter THC:CBD 1:1 ebenfalls deutlich über denen der jeweiligen Vergleichsgruppe: GWMS0001 26,1 % (Placebo 11,1 %), GWMS0106 17,5 % (Placebo 9,4 %), GWCL0403 13 % (Placebo 11 %) und GWSP0604 45 % (Placebo 33 %). Die Odds ratio der Studien GWMS0001, GWMS0106, GWCL0403 und GWSP0604 betrug 3,04 (95%-CI: 0,51; 18,11), 2,05 (95 %-CI: 0,78; 5,37), 1,21 (95 %-CI: 0,62; 2,37) bzw. 1,65 (95 %-CI: 0,98; 2,78), war statistisch aber in keiner Studie signifikant.

Der Grund für die deutlich höhere Anzahl von Patienten, deren Symptomatik sich in der Studie GWSP0604 klinisch relevant um 30 % bzw. 50 % verbesserte, liegt daran, dass zur

Zielpopulation schwerstkranke Patienten zählen. Diese profitieren schon wesentlich von einer Verbesserung von 30 %. Da die Fallzahl der Patienten mit einer Verbesserung von mindestens 50 % sehr klein ist, ist eine statistische Signifikanz nicht zu erwarten gewesen.

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.

Die Studie GWSP0604 ist die pivotale Zulassungsstudie. Die Studien GWMS0001, GWMS0106 und GWCL0403 haben supportiven Charakter und überprüfen die klinische Relevanz in der Gruppe der Responder.

Es wurde für den Endpunkt „Klinisch relevanter Rückgang der Schwere der Spastizität um mindestens 30 %“ für folgende Studien eine Meta-Analyse mit dem Programm RevMan 5.1 durchgeführt: GWMS0001, GWMS0106 und GWCL0403. In diesen drei Studien wurde der Endpunkt in einem vergleichbaren Design bei Patienten, die bezüglich einer Vorbehandlung mit THC:CBD 1:1 naive waren, analysiert.

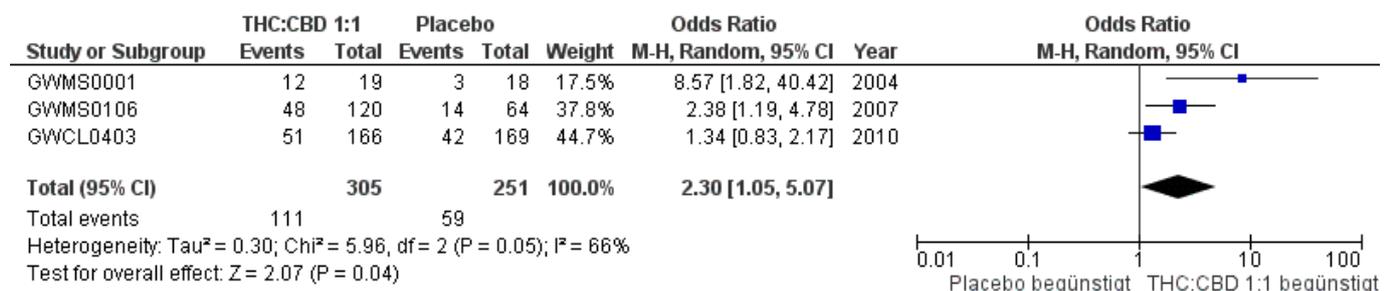


Abbildung 5: Überprüfung der Homogenität für den Endpunkt „Klinisch relevanter Rückgang der Schwere der Spastizität“ (Responder, mind. 30 %)

Wie aus Abbildung 5 ersichtlich ist, können die drei Studien nicht zu einer Meta-Analyse zusammengefasst werden, weil die Heterogenität zwischen den Studien stark ausgeprägt ist, wahrscheinlich hervorgerufen durch den geringen Studienpool, der in GWMS0001 analysiert wurde bzw. durch die hohe Abweichung der ITT- von der PP-Population in der Studie GWCL0403. Allerdings ist tendenziell ein relevanter Rückgang der Schwere der Spastizität zu beobachten (OR 2,30 und p= 0,04).

Es wurde eine Sensitivitätsanalyse nur mit den beiden Studien GWMS0106 und GWCL0403 durchgeführt, wobei die Heterogenität hier auch 44 % betrug. In dieser Analyse lag die Responderrate (mindestens 30 %ige Verbesserung) unter der Wirkstoffkombination THC:CBD 1:1 bei 34,6 % gegenüber der Vergleichsgruppe mit 24 % (p = 0,06). Die Odds ratio betrug 1,69 zu Gunsten der Wirkstoffkombination THC:CBD 1:1; diese Unterschiede waren nicht signifikant.

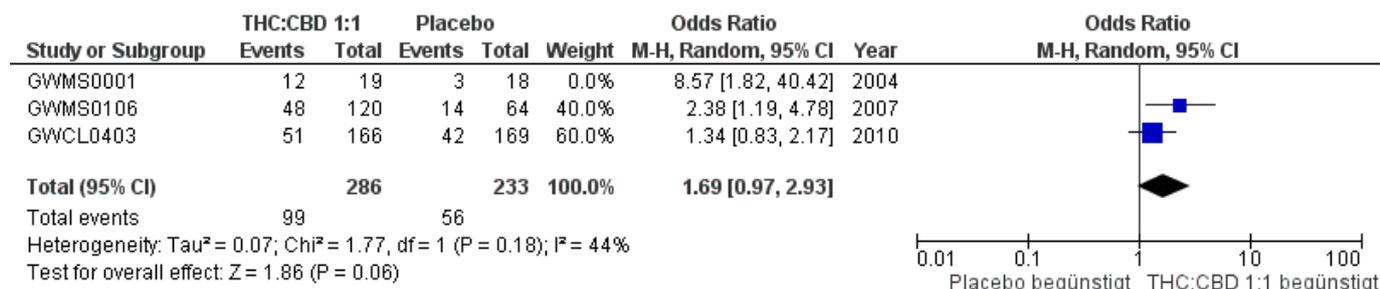


Abbildung 6: Meta-Analyse für den Endpunkt „Klinisch relevanter Rückgang der Schwere der Spastizität“ (Responder, mind. 30%) ohne die Studie GWMS0001

In einer vor kurzem veröffentlichten Meta-Analyse zeigt Wade et al. in einer Responder-Analyse eine Odds ratio von 1,62 (3). Allerdings berücksichtigte diese Analyse bei der Studie GWMS0001 alle Patienten mit Spastik, ging also von einer anderen Grundgesamtheit aus, die in der Publikation nicht weiter erläutert wurde. Da 97,5 % der THC:CBD 1:1 und 93,8 % der Placebo-Gruppe in dieser Studie anti-spastische Medikation erhielt, wird davon ausgegangen, dass alle Patienten unter mittelschwerer bis schwerer Spastik litten.

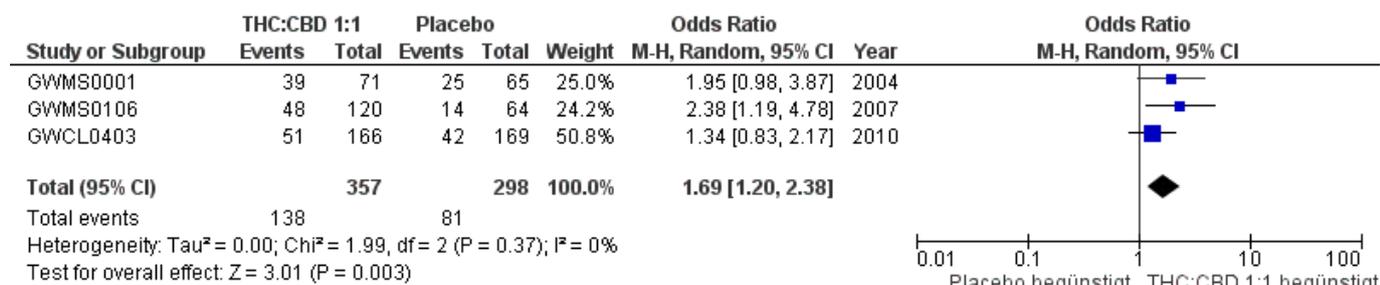


Abbildung 7: Meta-Analyse für den Endpunkt „Klinisch relevanter Rückgang der Schwere der Spastizität“ (Responder, mind. 30%) inklusive aller Patienten mit Spastik aus GWMS0001

Dieselben Daten wurden auch im Rahmen dieser Nutzenbewertung einer Analyse unterzogen (Abbildung 7), dabei aber nur die Patienten der ITT-Population laut Clinical Study Report berücksichtigt. In diesem Fall zeigt sich Homogenität zwischen den Studien und ein statistisch signifikanter Vorteil der THC:CBD 1:1 Add-on Therapie gegenüber Placebo (p=0,003), der in einer Odds ratio von 1,69 resultiert. Es befanden sich 38,7 % der THC:CBD 1:1 Patienten und 27,2 % der Placebo-Patienten unter den Respondern.

Der RCT der Studie GWSP0604 wurde nicht in eine erweiterte Meta-Analyse integriert, da dieses RCT eine andere Patientenpopulation untersuchte. Es wurde die Wirksamkeit von Sativex[®] bei Patienten untersucht, die auf einen Therapieversuch mit Sativex[®] bereits positiv angesprochen hatten.

Im Rahmen der Responder-Analyse wurde eine Sensitivitätsanalyse des Endpunktes „Klinisch relevanter Rückgang der Schwere der Spastizität um 50%“ durchgeführt. Es zeigt

sich ein numerischer, jedoch statistisch nicht signifikanter Vorteil der Add-on Therapie gegenüber dem Placebo ($p=0,1$) in den drei Studien (Abbildung 8).

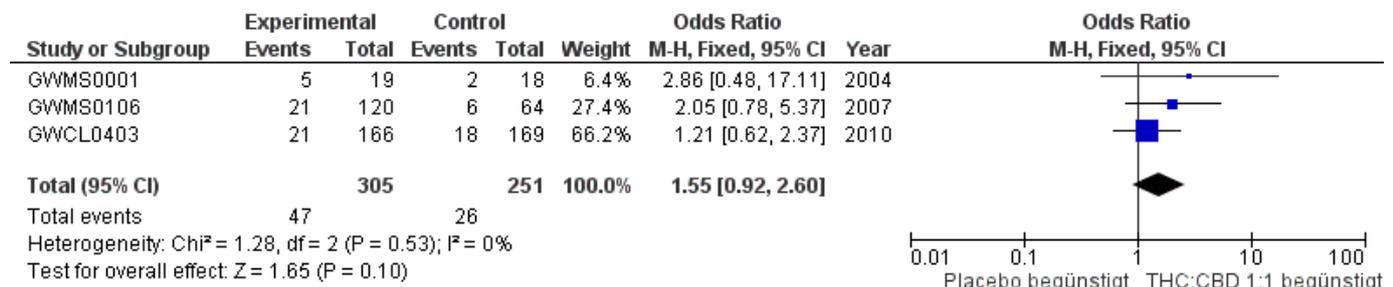


Abbildung 8: Meta-Analyse für den Endpunkt „Klinisch relevanter Rückgang der Schwere der Spastizität“ (Responder, mind. 50 %) inklusive aller Patienten mit Spastik aus GWMS0001

Die Ergebnisse der Meta-Analysen unterstützen die Ergebnisse aus der pivotalen Zulassungsstudie GWSP0604 und zeigen einen Vorteil der Behandlung von THC:CBD 1:1 gegenüber Placebo bezüglich des Endpunktes „Klinisch relevanter Rückgang der Schwere der Spastizität“.

4.3.1.3.1.3 Anzahl Spastiken pro Tag – 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.

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

Tabelle 4-20: Operationalisierung von Anzahl Spastiken pro Tag

Studie	Operationalisierung
GWMS0106	Die Anzahl der Spastiken pro Tag wurden von den Patienten im Tagebuch täglich über die gesamte Studiendauer von 8 Wochen dokumentiert (2 Wochen Screeningphase plus 6 Wochen Behandlungsphase) oder bis zum vorzeitigen Studienende. Die Anzahl der Spastiken pro Tag (Durchschnittswert bei der Betrachtung einer Woche) wurde klassifiziert erfasst: 0 = keine Spastik, 1 = 1 oder weniger Spastiken pro Tag, 2 = Zwischen 1 und 5 Spastiken pro Tag, 3 = 6 bis 9 Spastiken pro Tag, 4 = 10 oder mehr Spastiken pro Tag oder kontinuierliche Spastik. Zur Bestimmung der Änderung der Anzahl der Spastiken pro Tag wurden die Werte vom Studienende mit den Baseline-Werten verglichen.
GWSP0604	Die absolute Anzahl der Spastiken pro Tag wurden von den Patienten im Tagebuch täglich über die gesamte Studiendauer dokumentiert (1 Woche Screeningphase plus 4 Wochen Studienmedikation Run-in Phase plus 12 Wochen randomisierte Fortsetzung versus Absetzen der Studienmedikation bei Ansprechen auf die Studienmedikation in der Run-in Phase oder bis zum vorzeitigen Studienende). Zur Bestimmung der Änderung der Anzahl der Spastiken pro Tag wurden die Werte vom Studienende mit den Baseline-Werten verglichen.

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 4-21: Bewertung des Verzerrungspotenzials für Anzahl Spastiken pro Tag 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
GWMS0106	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWSP0604	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Aufgrund des niedrigen Verzerrungspotentials auf Studienebene aller relevanten RCT, die den Endpunkt „Anzahl der Spastiken pro Tag“ untersucht haben, kann von einem niedrigen Verzerrungspotential auch auf Ebene des Endpunkts selbst ausgegangen werden. In den beiden Studien waren die Endpunkterheber (Patienten) verblindet. Der Endpunkt wurde wie geplant für die jeweilige ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt wie geplant statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in beiden Studien das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

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

Tabelle 4-22: Ergebnisse für Anzahl Spastiken pro Tag aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anzahl Spastiken pro Tag(nach Häufigkeitsklassen) (Klassiert: 0 = keine Spastik, 1 = eine oder weniger Spastiken pro Tag, 2 = zwischen einer und fünf Spastiken pro Tag, 3 = sechs bis neun Spastiken pro Tag, 4 = zehn oder mehr Spastiken pro Tag oder dauerhafte Spastizität)								
	Baseline			Studienende			Veränderung		
	N	Mean (SD)	Median Min-Max	N	Mean (SD)	Median Min-Max	Mean (SD)	Median Min-Max	Gruppen-differenz (*) (95%-CI) p-Wert
GWMS 0106 ¹⁾									
Sativex®	123	2,27 (1,13)	2,29 (0,0-4,0)	121	1,92 (1,13)	2,0 (0,0-4,0)	-0,37 (0,76)	-0,26 (-3,4-1,4)	-0,17 (-0,39, 0,06)
Placebo	64	2,54 (1,12)	2,62 (0,0-4,0)	65	2,27 (1,16)	2,0 (0,0-4,0)	-0,26 (0,78)	-0,05 (-2,8-1,9)	p=0,141
Durchschnittliche Anzahl Spastiken pro Tag									
	Baseline			Studienende			Veränderung		
	N	Mean (SD)	Median Min-Max	N	Mean (SD)	Median Min-Max	Mean (SD)	Median Min-Max	Gruppen-differenz (*) (95%-CI) p-Wert
GWSP 0604 ²⁾									
Sativex®	124	5,61 (9,09)	3,14 (0,0-74,1)	124	5,56 (10,33)	2,86 (0,0-89,0)	-0,05 (4,82)	0,0 (-25,1-19,9)	-2,53 (-4,27, -0,79)
Placebo	117	5,29 (7,3)	3,14 (0,0-42,9)	117	7,70 (12,02)	3,86 (0,0-77,9)	2,41 (8,55)	0,29 (-19,0-73,1)	p=0,004
(*) Adjustiertes ANCOVA Modell									
1. Verbesserung 6 Wochen nach Behandlungsbeginn									
2. Verbesserung 12 Wochen nach Fortführung versus Absetzen nach Ansprechen (mindestens 20 % Verbesserung der NRS) auf 4 Wochen Run-in Behandlung mit Sativex®									

Die Anzahl der Spastiken pro Tag wurde in beiden RCT vom Patienten erfasst und ausgewertet. Im direkten Vergleich zur Therapiefortsetzung ohne die Wirkstoffkombination THC:CBD 1:1 zeigt die Add-on Therapie unter der Wirkstoffkombination THC:CBD 1:1 in beiden RCT eine niedrigere Anzahl an Spasmen pro Tag. Bei Patienten, die auf die Therapie mit THC:CBD 1:1 angesprochen haben (GWSP0604), betrug die durchschnittliche absolute Zahl Spasmen nach Fortsetzung der Add-on Therapie mit THC:CBD 1:1 pro Tag 5,56 (SD:

10,33) (Placebo 7,70 (SD 12,02)). Die Gruppendifferenz bei diesen Patienten betrug -2,53 Anfälle pro Tag (95 %-CI: -4,27, -0,79) zugunsten der Wirkstoffkombination THC:CBD 1:1.

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.

Eine Meta-Analyse wurde nicht durchgeführt, da in den beiden Studien, in denen dieser Endpunkt erhoben wurde, unterschiedliche Patientenpopulationen in einem unterschiedlichen Behandlungsmuster betrachtet wurden. In Studie GWMS0106 wurden naive Patienten im Follow-up mit der Wirkstoffkombination THC:CBD 1:1 oder Placebo behandelt, und in Responderstudie GWSP0604 wurde das Fortführen von THC:CBD 1:1 versus Absetzen (Placebogruppe) bei Patienten untersucht, die zuvor bereits auf eine 4-wöchige THC:CBD 1:1-Behandlung positiv angesprochen hatten.

4.3.1.3.1.4 Schlafqualität – 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.

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 Schlafqualität

Studie	Operationalisierung
GWCL0403	Die Schlafqualität wurde über die gesamte Studiendauer an den Visiten Woche 0, 6, 10 und 14 oder bis zum vorzeitigen Studienende erhoben. Die Dokumentation der Schlafqualität erfolgte auf einer 11-stufigen NRS (0= 'Slept extremely well' - 10= 'Unable to sleep at all'). Die Auswertung erfolgte als Differenz der Beurteilung im Vergleich Studienende zu Baseline (Randomisierung Woche 0).
GWSP0604	Die Schlafqualität wurde als Unterbrechung des Schlafs von den Patienten im Tagebuch täglich über die gesamte Studiendauer dokumentiert (1 Woche Screeningphase plus 4 Wochen Studienmedikation Run-in Phase plus 12 Wochen randomisierte Fortsetzung versus Absetzen der Studienmedikation bei Ansprechen auf die Studienmedikation in der Run-in Phase oder bis zum vorzeitigen Studienende). Die Dokumentation der Schlafqualität erfolgte auf einer 11-stufigen NRS (0= 'No Sleep Disruption' - 10= 'Worst Possible Sleep Disruption'). Die Auswertung erfolgte als Differenz der Beurteilung im Vergleich Studienende zu Baseline (bei Randomisierung in Woche 4/5).

Für die Studie GWMS0001 wurde die Schlafqualität untersucht. Allerdings erfolgte bei der Auswertung keine Subgruppenanalyse für die 37 Patienten mit primärer Spastik. Aus diesem Grund wurde diese Studie für den Endpunkt Schlafqualität nicht berücksichtigt.

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 Schlafqualität 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
GWCL0403	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWSP0604	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Aufgrund des niedrigen Verzerrungspotentials auf Studienebene aller relevanten RCT, die den Endpunkt Schlafqualität auf der 11-stufigen NRS untersucht haben, kann von einem niedrigen Verzerrungspotential auch auf Ebene des Endpunkts selbst ausgegangen werden. In beiden Studien waren die Endpunkterheber (Patienten) verblindet. Der Endpunkt wurde, wie geplant, für die jeweilige ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt wie geplant statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in allen drei Studien das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

Stellen Sie die Ergebnisse für den Endpunkt Schlafqualität 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 Schlafqualität aus RCT mit dem zu bewertenden Arzneimittel

Studie	Schlafqualität (NRS 0-10 Beurteilung) Erhoben als Schlafqualität (GWCL0403) oder als Schlafunterbrechung (GWSP0604)								
	Baseline			Studienende			Veränderung		
	N	Mean (SD)	Median Min-Max	N	Mean (SD)	Median Min-Max	Mean (SD)	Median Min-Max	Gruppen- differenz (*) (95%-CI) p-Wert
GWCL 0403 ¹⁾									
Sativex®	166	3,3 (2,95)	3,0 (0-10)	149	2,4 (2,52)	2,0 (0-8)	-0,8 (2,83)	0,0 (-9-8)	-0,07 (-0,55, 0,4) p=0,7637
Placebo	169	3,0 (2,8)	3,0 (0-10)	158	2,3 (2,51)	2,0 (0-10)	-0,6 (2,38)	0,0 (-8, 6)	
GWSP 0604 ²⁾									
Sativex®	124	1,96 (1,80)	1,5 (0,0-6,7)	124	1,71 (1,65)	1,21 (0,0-6,3)	-0,25 (1,06)	0,0 (-3,9-3,9)	-0,88 (-1,25, - 0,51) p<0,001
Placebo	117	2,07 (1,75)	1,86 (0,0-7,0)	117	2,65 (2,45)	2,0 (0,0-9,0)	0,59 (1,80)	0,29 (-3,9-7,0)	
(*) Adjustiertes ANCOVA Modell									
1. Verbesserung 14 Wochen nach Behandlungsbeginn									
2. Verbesserung 12 Wochen nach Fortführung versus Absetzen nach Ansprechen (mindestens 20% Verbesserung der NRS) auf 4 Wochen Run-in Behandlung mit Sativex®									

Im direkten Vergleich zur Therapiefortsetzung ohne THC:CBD 1:1 zeigt die Add-on Therapie unter THC:CBD 1:1 in beiden RCT, die den Endpunkt Schlafqualität untersuchten, eine Besserung der Schlafqualität. In der Studie GWCL0403 betrug die Differenz im Durchschnitt -0,07 (95 %-CI: -0,55; 0,4) auf der 11-stufigen NRS im Vergleich zu Placebo. Bei der Population, die dem Zulassungsstatus von Sativex® am besten entspricht, also den Patienten, die auf eine Therapie mit Sativex® in klinisch relevanter Form ansprechen (Responderstudie GWSP0604), betrug die Differenz im Durchschnitt -0,88 (95 %-CI: -1,25; -0,51) und war statistisch signifikant (p < 0,001). Hier zeigte sich die Verbesserung der Schlafqualität in einer Abnahme der Schlafunterbrechungen in Relation zur Vergleichstherapie.

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.

Eine Meta-Analyse wurde nicht durchgeführt, da in den beiden Studien, in denen dieser Endpunkt erhoben wurde, unterschiedliche Patientenpopulationen in einem unterschiedlichen Behandlungsmuster betrachtet wurden. In Studie GWCL0403 wurden naive Patienten im Follow-up mit THC:CBD 1:1 oder Placebo behandelt, in der Responderstudie GWSP0604 wurde das Fortführen von THC:CBD 1:1 versus Absetzen (Placebogruppe) bei Patienten untersucht, die zuvor bereits auf eine 4-wöchige THC:CBD 1:1 Behandlung positiv angesprochen hatten. Zudem wurde der Endpunkt in den einzelnen Studien mit unterschiedlichen Fragestellungen als Schlafqualität oder Schlafunterbrechung erhoben.

4.3.1.3.1.5 Gesundheitszustand (EQ-5D) – 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.

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 Gesundheitszustand (EQ-5D)

Studie	Operationalisierung
GWCL0403	Der Gesundheitszustand wurde vom Patienten an Baseline (Woche 0) und am Studienende (Woche 14 oder bei vorzeitigem Studienende) dokumentiert. Die Erfassung erfolgte mittels der validierten Fragebögen EQ-5D.
GWSP0604	Der Gesundheitszustand wurde vom Patienten an Behandlungsbeginn (Woche 0), zum Zeitpunkt der randomisierten Fortsetzung oder Abbruch (Woche 4/5) und am Studienende (Woche 18 oder bei vorzeitigem Studienende) dokumentiert. Die Erfassung erfolgte mittels der validierten Fragebögen EQ-5D.

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 den Gesundheitszustand(EQ-5D) 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
GWCL0403	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWSP0604	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Aufgrund des niedrigen Verzerrungspotentials auf Studienebene aller relevanten RCT, die den Endpunkt Gesundheitszustand untersucht haben, kann von einem niedrigen Verzerrungspotential auch auf Ebene des Endpunkts selbst ausgegangen werden. In den beiden Studien waren die Endpunkterheber (Patienten) verblindet. Der Endpunkt wurde, wie geplant, für die jeweilige ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt, wie geplant, statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in allen beiden Studien das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

Stellen Sie die Ergebnisse für den Endpunkt Gesundheitszustand 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 Gesundheitszustand (EQ-5D) aus RCT mit dem zu bewertenden Arzneimittel

Studie	Gesundheitszustand EQ-5D VAS Beurteilung								
	Baseline			Studienende			Veränderung		
	N	Mean (SD)	Median Min-Max	N	Mean (SD)	Median Min-Max	Mean (SD)	Median Min-Max	Gruppen- differenz (*) (95%-CI) p-Wert
GWCL 0403 ¹⁾									
Sativex®	165	51,8 (19,01)	52,0 (0-98)	159	56,6 (18,72)	59,0 (3-95)	4,1 (19,25)	5,0 (-50-55)	1,42 (-3,11;5,94) p=0,5376
Placebo	167	53,0 (18,35)	50,0 (0-100)	164	55,8 (18,62)	51,0 (15-100)	2,7 (22,31)	2,5 (-60-70)	
GWSP 0604 ²⁾									
Sativex®	123	59,1 (15,6)	60,0 (27-95)	121	58,3 (17,6)	60,0 (20-100)	-0,7 (15,3)	0,0 (-40-40)	1,24 (-3,01;-5,5) p=0,5644
Placebo	113	61,2 (16,0)	65,0 (10-91)	117	58,4 (19,3)	60,0 (10-97)	-2,8 (19,9)	0,0 (-64-50)	
(*) Adjustiertes ANCOVA Modell									
1. Verbesserung 14 Wochen nach Behandlungsbeginn									
2. Verbesserung 12 Wochen nach Fortführung versus Absetzen nach Ansprechen (mindestens 20 % Verbesserung der NRS) auf 4 Wochen Run-in Behandlung mit Sativex®									

Der Gesundheitszustand mittels EQ-5D wurde in zwei Studien erhoben. Im direkten Vergleich zur Therapiefortsetzung ohne THC:CBD 1:1 berichteten Patienten unter Zusatztherapie mit THC:CBD 1:1 eine im Gruppenvergleich um 1,42 bzw. 1,24 verbesserte Beurteilung des Gesundheitszustands im Behandlungszeitraum. Die resultierende Gruppendifferenz war in keiner der Studien statistisch signifikant.

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

Eine Meta-Analyse wurde nicht durchgeführt, da in den beiden Studien, in denen dieser Endpunkt erhoben wurde, unterschiedliche Patientenpopulationen in einem unterschiedlichen Behandlungsmuster betrachtet wurden. In Studie GWCL0403 wurden naive Patienten im Follow-up mit THC:CBD 1:1 oder Placebo behandelt. In der Responderstudie GWSP0604 wurde das Fortführen von THC:CBD 1:1 versus Absetzen (Placebogruppe) bei Patienten untersucht, die zuvor bereits auf eine 4-wöchige Sativex[®] Behandlung positiv angesprochen hatten.

4.3.1.3.1.6 Lebensqualität – 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.

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

Tabelle 4-29: Operationalisierung von Lebensqualität

Studie	Operationalisierung
GWCL0403	Die Erfassung der Lebensqualität erfolgte anhand der MSQoL-54 am Studienanfang und Studienende.
GWSP0604	Die Erfassung der Lebensqualität erfolgte anhand der SF-36 zur Randomisierung und zum Studienende.
ARA2009	Die Erfassung der Lebensqualität erfolgte anhand des VAS QoL zum Studienanfang und Studienende.

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 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
GWCL0403	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWSP0604	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ARA2009	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Aufgrund des niedrigen Verzerrungspotentials auf Studienebene aller relevanten RCT, die den Endpunkt Lebensqualität untersucht haben, kann von einem niedrigen Verzerrungspotential auch auf Ebene des Endpunkts selbst ausgegangen werden. In den drei Studien waren die Endpunkterheber (Patienten) verblindet. Der Endpunkt wurde wie geplant für die jeweilige ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt wie geplant statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in allen drei Studien das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

Stellen Sie die Ergebnisse für den Endpunkt Lebensqualität 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 Lebensqualität aus RCT mit dem zu bewertenden Arzneimittel*

Studie	Beurteilung der Lebensqualität					
	THC:CBD 1:1		Placebo		Gruppendifferenz [95%-CI]	p-Wert
	N	Mean	N	Mean		
GWCL0403						
Gesundheit	60	3,75	59	1,89	1,86 [-3,88; 7,59]	0,5230
Funktionsbeeinträchtigungen (gesundheitliche Probleme)	58	12,37	58	14,97	-2,60 [-15,36; 10,16]	0,6869
Funktionsbeeinträchtigungen (emotionale Probleme)	58	-5,08	57	-2,34	-2,74 [-19,60; 14,12]	0,7480
Schmerz	59	11,74	58	8,82	2,92 [-5,34; 11,18]	0,4856
Emotionales Wohlbefinden	60	-1,58	59	-0,53	-1,05 [-6,66; 4,55]	0,7106
Energie	60	2,22	59	5,95	-3,73 [-10,45; 3,08]	0,2806
Gesundheitswahrnehmung	60	0,72	59	5,61	-4,89 (-11,22; 1,44]	0,1291
Gesellschaftliche Funktion	60	3,18	58	6,64	-0,46 [-8,56; 7,63]	0,9101
Gesundheitliche Leiden	60	8,10	58	8,14	-0,05 [-8,37; 8,27]	0,9910
Kognitive Funktion	60	-1,27	59	8,96	-7,23 [-14,32; -0,13]	0,0460
Sexuelle Funktion	47	5,85	51	0,48	5,36 [-5,12; 15,85]	0,3126
Veränderung der Gesundheit	60	2,40	59	11,74	-9,35 [-17,04; -1,65]	0,0177
Zufriedenheit mit der sexuellen Funktion	47	2,62	50	5,67	-0,06 [-10,74; 10,63]	0,9917
Allgemeine Lebensqualität	60	2,63	59	6,05	-3,41 [-9,31; 2,48]	0,2534
Zusammenfassung der physischen Werte	46	5,10	50	6,61	-1,51 [-6,49; 3,47]	0,5494
Zusammenfassung der mentalenen Werte	58	-0,05	56	3,04	-3,09 [-9,14; 2,95]	0,3124

GWSP0604	THC:CBD 1:1		Placebo		Gruppendifferenz [95%-CI]	p-Wert
Gesundheitliche Funktion	121	0,30	117	0,76	-0,46 [-3,73; 2,81]	0,7819
Gesundheitliche Rolle	119	-0,31	116	0,98	-1,30 [-7,06; 4,47]	0,6582
Körperschmerz	118	-0,05	115	-5,06	5,01 [-0,21; 10,23]	0,0596
Allgemeine Gesundheit	117	1,20	114	-0,12	1,32 [-2,06; 4,70]	0,4424
Vitalität	119	-1,17	112	-3,35	2,19 [-2,01; 6,38]	0,3057
Gesellschaftliche Funktion	112	-0,97	103	-0,32	-0,65 [-6,97; 5,67]	0,8403
Emotionale Rolle	119	-1,26	115	1,53	-2,78 [-8,56; 2,99]	0,3431
Mentale Gesundheit	119	-2,20	112	-2,94	0,74 [-2,84; 4,33]	0,6826
ARA2009	Baseline-Wert (SD)		Post-Sativex-Wert (SD)		Post-Placebo-Wert (SD)	p-Wert
n=17	4,29 (2,08)		4,00 (2,00)		3,65 (2,29)	0,31

* Die „Mean“-Werte stellen die durchschnittlichen Differenzen der jeweiligen Behandlungsarme zwischen Baseline und Studienende dar.

Für die Studien GWCL0403 und GWSP0604 gilt: Ist die Gruppendifferenz positiv, ist eine Verbesserung zugunsten von THC:CBD 1:1 indiziert. Für die Studie ARA2009 ist auf einer Skala von 0-10 ein hoher Wert ein Zeichen für höhere Lebensqualität.

Die Einzelergebnisse zeigen in verschiedene Richtungen und lassen keine einheitlichen Aussagen zur Lebensqualität zu (Tabelle 4-31).

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

Eine Meta-Analyse wurde nicht durchgeführt, da in allen drei Studien unterschiedliche Bewertungsmethoden angewendet wurden, die untereinander nicht vergleichbar sind.

4.3.1.3.1.7 Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC)– 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.

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

Tabelle 4-32: Operationalisierung von Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC)

Studie	Operationalisierung
GWMS0106	Die Gesamtbeurteilung wurde vom Patienten auf Subject's Global impression of Change (SGIC) Skalen am Studienende (Woche 8 oder bei vorzeitigem Studienabbruch) dokumentiert. Der SGIC wurde erfasst als Beurteilung der Verbesserung des Gesamtzustands auf 7-Punkt Skalen (1 = <i>Very Much Improved</i> , 2 = <i>Much Improved</i> , 3 = <i>Minimally Improved</i> , 4 = <i>No Change</i> , 5 = <i>Minimally Worse</i> , 6 = <i>Much Worse</i> , 7 = <i>Very Much Worse</i>).
GWSP0604	Die Gesamtbeurteilung wurde auf Global Impression of Change Skalen im Studienverlauf und -ende (Woche 4/5 (Randomisierung), Woche 10, 14 und 18 oder bei vorzeitigem Studienende) vom Patienten (SGIC), Caregiver (CGIC) und Arzt (PhGIC) dokumentiert. Die GIC wurde erfasst als Beurteilung der Verbesserung des Gesamtzustands auf 7-Punkt Skalen (1 = <i>Very Much Improved</i> , 2 = <i>Much Improved</i> , 3 = <i>Minimally Improved</i> , 4 = <i>No Change</i> , 5 = <i>Minimally Worse</i> , 6 = <i>Much Worse</i> , 7 = <i>Very Much Worse</i>). Zur Ermittlung der Veränderung wurden die Werte zum Studienende mit den Werten verglichen, die zur Randomisierung erhoben wurden,.

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 4-33: Bewertung des Verzerrungspotenzials für Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC) 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
GWMS0106	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWSP0604	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Aufgrund des niedrigen Verzerrungspotentials auf Studienebene aller relevanten RCT, die den Endpunkt Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC) untersucht haben, kann von einem niedrigen Verzerrungspotential auch auf Ebene des Endpunkts selbst ausgegangen werden. In den beiden Studien waren die Endpunkterheber (Patienten) verblindet. Der Endpunkt wurde, wie geplant, für die jeweilige ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der

Endpunkt, wie geplant, statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in beiden Studien das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

Stellen Sie die Ergebnisse für den Endpunkt Globalbeurteilung der klinischen Veränderung durch den Patienten für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-34: Ergebnisse für Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC) aus RCT mit dem zu bewertenden Arzneimittel

Studie	Beurteilung der klinischen Änderung am Studienende durch den Patienten								
	Anzahl Patienten n (%)								
	N	Sehr viel besser	Viel besser	Wenig besser	Keine Änderung	Wenig schlechter	Viel schlechter	Sehr viel schlechter	Odds ratio (95%-CI) p-Wert
GWMS 0106¹⁾									
Sativex®	124	2 (2)	24 (19)	40 (32)	38 (31)	10 (8)	2 (2)	0 (0)	1,36 (0,78, 2,37) p=0,279
Placebo	65	0 (0)	11 (17)	20 (31)	26 (40)	5 (8)	2 (3)	0 (0)	
GWSP 0604²⁾									
Sativex®	121	14 (12)	40 (33)	39 (32)	17 (14)	4 (3)	6 (5)	1 (1)	1,703 (1,075, 2,698) p=0,0234
Placebo	117	10 (9)	28 (24)	33 (28)	35 (30)	7 (6)	1 (1)	3 (3)	
1. Verbesserung 6 Wochen nach Behandlungsbeginn 2. Verbesserung 12 Wochen nach Fortführung versus Absetzen nach Ansprechen (mindestens 20 % Verbesserung der NRS) auf 4 Wochen Run-in Behandlung mit Sativex®									

Die Globalbeurteilung der Veränderung der Spastik durch den Patienten (SGIC) wurde in zwei Studien erhoben. Im direkten Vergleich zur Therapiefortsetzung ohne THC:CBD 1:1 zeigt die Therapiefortsetzung unter der Wirkstoffkombination THC:CBD 1:1 häufiger eine bessere abschließende Beurteilung der globalen Veränderung durch die Patienten. Die resultierende Odds ratio betrug in Studie GWMS0106 1,36 (95 %-CI: 0,78, 2,37) und in Studie GWSP0604 1,703 (95 %-CI: 1,075, 2,698) jeweils zugunsten von THC:CBD 1:1. Zu beachten ist dabei, dass in der Responderstudie GWSP0604, die dem Zulassungsstatus von Sativex® am besten entspricht, d.h. bei Patienten, die auf eine Therapie mit THC:CBD 1:1 in klinisch relevanter Form ansprechen, die Änderung statistisch signifikant ist.

Die Odds ratio wurde mittels der logistischen Regressionstechnik berechnet.

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.

Eine Meta-Analyse wurde nicht durchgeführt, da in beiden Studien, in denen dieser Endpunkt erhoben wurde, unterschiedliche Patientenpopulationen in einem unterschiedlichen Behandlungsmuster betrachtet wurden. In Studie GWMS0106 wurden naive Patienten im Follow-up mit Sativex[®] oder Placebo behandelt, in der Responderstudie GWSP0604 wurde das Fortführen von Sativex[®] versus Absetzen (Placebogruppe) bei Patienten untersucht, die zuvor bereits auf eine 4-wöchige THC:CBD 1:1-Behandlung positiv angesprochen hatten.

4.3.1.3.1.8 Globalbeurteilung der klinischen Veränderung durch die betreuende Person (CGIC) – 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.

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

Tabelle 4-35: Operationalisierung von Globalbeurteilung der klinischen Veränderung durch die betreuende Person (CGIC)

Studie	Operationalisierung
GWCL0403	Die Gesamtbeurteilung wurde vom Caregiver auf Caregiver's Global impression of Change (CGIC) Skalen am Studienende (Woche 14 oder bei vorzeitigem Studienabbruch) dokumentiert. Der CGIC wurde erfasst als Beurteilung der Verbesserung des Gesamtzustands auf 7-Punkt Skalen (1 = <i>Very Much Improved</i> , 2 = <i>Much Improved</i> , 3 = <i>Minimally Improved</i> , 4 = <i>No Change</i> , 5 = <i>Minimally Worse</i> , 6 = <i>Much Worse</i> , 7 = <i>Very Much Worse</i>).
GWSP0604	Die Gesamtbeurteilung wurde auf Global Impression of Change Skalen im Studienverlauf und -ende (Woche 4/5 (Randomisierung), Woche 10, 14 und 18 oder bei vorzeitigem Studienende) vom Patienten (SGIC), Caregiver (CGIC) und Arzt (PhGIC) dokumentiert. Die GIC wurde erfasst als Beurteilung der Verbesserung des Gesamtzustands auf 7-Punkt Skalen (1 = <i>Very Much Improved</i> , 2 = <i>Much Improved</i> , 3 = <i>Minimally Improved</i> , 4 = <i>No Change</i> , 5 = <i>Minimally Worse</i> , 6 = <i>Much Worse</i> , 7 = <i>Very Much Worse</i>). Zusätzlich zur allgemeinen funktionellen Globalbeurteilung der klinischen Veränderung wurde von den Caregivern die Änderung der Transferierbarkeit auf der GIC bestimmt. Zur Einschätzung der Veränderung wurden die Werte zum Studienende mit den Werten verglichen, die zur Randomisierung erhoben wurden.

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 4-36: Bewertung des Verzerrungspotenzials für Globalbeurteilung der klinischen Veränderung durch die betreuende Person (CGIC) 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
GWCL0403	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWSP0604	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Aufgrund des niedrigen Verzerrungspotentials auf Studienebene aller relevanten RCT, die den Endpunkt Globalbeurteilung der klinischen Veränderung durch die betreuende Person (CGIC) untersucht haben, kann von einem niedrigen Verzerrungspotential auch auf Ebene des Endpunkts selbst ausgegangen werden. In den beiden Studien waren die Endpunkterheber (Betreuende Person) verblindet. Der Endpunkt wurde, wie geplant, für die jeweilige ITT-

Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt, wie geplant, statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in beiden Studien das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

Stellen Sie die Ergebnisse für den Endpunkt Globalbeurteilung der klinischen Veränderung durch die betreuende Person für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-37: Ergebnisse für Globalbeurteilung der klinischen Veränderung durch die betreuende Person (CGIC) aus RCT mit dem zu bewertenden Arzneimittel

Studie	Beurteilung der klinischen Änderung am Studienende durch die betreuende Person								
	Funktionale Möglichkeiten des Patienten								
Anzahl Patienten n (%)									
	N	Sehr viel besser	Viel besser	Wenig besser	Keine Änderung	Wenig schlechter	Viel Schlechter	Sehr viel Schlechter	Odds ratio (95%-CI) p-Wert
GWCL 0403 ¹⁾									1,248
Sativex®	167	3 (2)	18 (13)	51 (36)	52 (37)	11 (8)	4 (3)	2 (1)	(0,0842; 1,849) p=0,2701
Placebo	170	5 (3)	14 (10)	37 (26)	67 (47)	16 (11)	4 (3)	1 (1)	
GWSP 0604 ²⁾									2,4
Sativex®	71	2 (3)	15 (21)	31 (44)	16 (23)	4 (6)	2 (3)	1 (1)	(1,29; 4,44) p=0,0053
Placebo	69	0	11 (16)	19 (28)	24 (35)	8 (12)	5 (7)	2 (3)	
Transferierbarkeit des Patienten									
Anzahl Patienten n (%)									
GWSP 0604 ²⁾									1,792
Sativex®	71	1 (1)	9 (13)	27 (38)	24 (34)	6 (8)	4 (6)	0	(0,973; 3,301) p=0,0613
Placebo	69	1 (1)	8 (12)	14 (20)	30 (43)	11 (16)	3 (4)	2 (3)	
1. Verbesserung 14 Wochen nach Behandlungsbeginn 2. Verbesserung 12 Wochen nach Fortführung versus Absetzen nach Ansprechen (mindestens 20% Verbesserung auf der NRS) auf 4 Wochen Run-in Behandlung mit Sativex®									

Die Globalbeurteilung der Veränderung der Spastik durch die betreuende Person (CGIC) wurde ebenfalls in zwei Studien erhoben. Dabei wurde in allen Studien nach der Beurteilung der Veränderung der funktionalen Möglichkeiten des Patienten gefragt, sowie in einer der beiden Studien (GWSP0604) zusätzlich nach der Beurteilung der Veränderung in der Transferierbarkeit der Patienten. Im direkten Vergleich zur Therapiefortsetzung ohne THC:CBD 1:1 (Placebo Gruppe) resultiert die Therapiefortsetzung unter THC:CBD 1:1 häufiger in einer besseren abschließenden Beurteilung der globalen Veränderung der funktionalen Möglichkeiten des Patienten durch die betreuenden Personen. Die resultierende Odds ratio betrug in Studie GWCL0403 1,248 (95 %-CI: 0,0842; 1,849) und in Studie GWSP0604 2,4 (95 %-CI: 1,29; 4,44), jeweils zugunsten von THC:CBD 1:1. Die abschließende Beurteilung der globalen Veränderung der Transferierbarkeit des Patienten ergab ebenfalls eine häufig bessere abschließende Beurteilung durch die betreuenden Personen bei Therapiefortsetzung unter THC:CBD 1:1 im direkten Vergleich zur Therapiefortsetzung ohne THC:CBD 1:1 (Placebo Gruppe). Die resultierende Odds ratio betrug in Studie GWSP0604 1,792 (95 %-CI: 0,973; 3,301) zugunsten von THC:CBD 1:1. Zu beachten ist, dass in der Responderstudie GWSP0604 der Effekt nach Fortsetzen (Sativex[®]-Gruppe) der Therapie für Patienten beurteilt wurde, die vorher positiv auf die Behandlung mit THC:CBD 1:1 angesprochen hatten.

Die Odds ratio wurde mittels der logistischen Regressionstechnik berechnet.

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.

Eine Meta-Analyse wurde nicht durchgeführt, da in den Studien unterschiedliche Patientenpopulationen in einem unterschiedlichen Behandlungsmuster betrachtet wurden. In Studie GWCL0403 wurden naive Patienten im Follow-up mit Sativex[®] oder Placebo behandelt, in der Responderstudie GWSP0604 wurde das Fortführen von Sativex[®] versus Absetzen (Placebogruppe) bei Patienten untersucht, die zuvor bereits auf eine 4-wöchige Sativex[®] Behandlung positiv angesprochen hatten.

4.3.1.3.1.9 Globalbeurteilung der klinischen Veränderung durch den Arzt (PhGIC) – 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.

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

Tabelle 4-38: Operationalisierung von Globalbeurteilung der klinischen Veränderung durch den Arzt (PhGIC)

Studie	Operationalisierung
GWSP0604	Die Gesamtbeurteilung wurde auf Global Impression of Change Skalen im Studienverlauf und -ende (Woche 4/5 (Randomisierung), Woche 10, 14 und 18 oder bei vorzeitigem Studienende) vom Patienten (SGIC), Caregiver (CGIC) und Arzt (PhGIC) dokumentiert. Die GIC wurde erfasst als Beurteilung der Verbesserung des Gesamtzustands auf 7-Punkt Skalen (1 = <i>Very Much Improved</i> , 2 = <i>Much Improved</i> , 3 = <i>Minimally Improved</i> , 4 = <i>No Change</i> , 5 = <i>Minimally Worse</i> , 6 = <i>Much Worse</i> , 7 = <i>Very Much Worse</i>). Es wurden die Werte zum Studienende mit den Werten, die bei der Randomisierung erhoben wurden, zur Einschätzung der Veränderung verglichen.

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 4-39: Bewertung des Verzerrungspotenzials für Globalbeurteilung der klinischen Veränderung durch den Arzt (PhGIC) 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
GWSP0604	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Aufgrund des niedrigen Verzerrungspotentials auf Studienebene der relevanten RCT, die den Endpunkt Globalbeurteilung der klinischen Veränderung durch den Arzt (PhGIC) untersucht hat, kann von einem niedrigen Verzerrungspotential auch auf Ebene des Endpunkts selbst ausgegangen werden. In der Studie waren die Endpunkterheber (Study Investigator) am Zeitpunkt der Endpunkterhebung Ende der doppelblinden randomisierten Studienphase B verblindet. Der Endpunkt wurde wie geplant für die ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt, wie geplant, statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in der Studie das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

Stellen Sie die Ergebnisse für den Endpunkt Globalbeurteilung der klinischen Veränderung durch den Arzt für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-40: Ergebnisse für Globalbeurteilung der klinischen Veränderung durch den Arzt (PhGIC) aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anzahl Patienten n (%)								
	Beurteilung der klinischen Änderung am Studienende durch den Arzt								
	N	Sehr viel besser	Viel besser	Wenig besser	Keine Änderung	Wenig schlechter	Viel Schlechter	Sehr viel Schlechter	Odds ratio (95%-CI) p-Wert
GWSP 0604 ¹⁾									
Sativex®	121	6 (5)	44 (36)	45 (37)	17 (14)	4 (3)	6 (5)	0	1,958
Placebo	115	8 (7)	29 (25)	26 (22)	37 (32)	13 (11)	4 (3)	0	(1,232, 3,112) p=0,0045
1. Verbesserung 12 Wochen nach Fortführung versus Absetzen nach Ansprechen (mindestens 20 % Verbesserung der NRS) auf 4 Wochen Run-in Behandlung mit Sativex®									

Die Globalbeurteilung der Veränderung der Spastik durch den Arzt (PhGIC) wurde in der Studie (GWSP0604) erhoben, d.h. der Patientenpopulation, die dem Zulassungsstatus von Sativex® am besten entspricht, also den Patienten, die auf eine Therapie mit Sativex® in klinisch relevanter Form ansprechen. Im direkten Vergleich zur Therapiefortsetzung ohne Sativex® zeigt die Therapiefortsetzung unter Sativex® häufiger eine bessere abschließende Beurteilung der globalen Veränderung des Patienten durch den behandelnden Arzt. Die resultierende Odds ratio betrug 1,958 (95 %-CI: 1,232, 3,112) zugunsten von Sativex®.

Die Odds ratio wurde mittels der logistischen Regressionstechnik berechnet.

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.

NA – Der Endpunkt wurde nur in einer RCT erhoben

4.3.1.3.1.10 Aktivitäten des täglichen Lebens (Barthel Index) – 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.

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

Tabelle 4-41: Operationalisierung von Aktivitäten des täglichen Lebens (Barthel Index)

Studie	Operationalisierung
GWCL0403	Die Fähigkeit zur Ausübung von Aktivitäten des täglichen Lebens wurde mittels des Barthel Index an Baseline (Woche 0) und am Studienende (Woche 14 oder früher bei vorzeitigem Studienende) vom Arzt dokumentiert.
GWSP0604	Die Fähigkeit zur Ausübung von Aktivitäten des täglichen Lebens wurde mittels des Barthel Index am Behandlungsbeginn (Woche 0), zum Zeitpunkt der randomisierten Fortsetzung oder Abbruch (Woche 4/5) und am Studienende (Woche 18 oder bei vorzeitigem Studienende) vom Arzt dokumentiert.

Für die Studie GWMS0001 wurde Aktivität des täglichen Lebens mit Hilfe des Barthel Index untersucht. Allerdings erfolgte dabei bei der Auswertung für die 37 Patienten mit primärer Spastik. Aus diesem Grund wurde diese Studie für den Endpunkt Aktivitäten des täglichen Lebens nicht berücksichtigt.

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 4-42: Bewertung des Verzerrungspotenzials für Aktivitäten des täglichen Lebens (Barthel Index) 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
GWCL0403	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWSP0604	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Aufgrund des niedrigen Verzerrungspotentials auf Studienebene aller relevanten RCT, die den Endpunkt Aktivitäten des täglichen Lebens (Barthel Index) untersucht haben, kann von einem niedrigen Verzerrungspotential auch auf Ebene des Endpunkts selbst ausgegangen werden. In den beiden Studien waren die Endpunkterheber (Study Investigator) am Zeitpunkt der Endpunkterhebung verblindet. Zwar waren in der ersten Phase in Studie GWSP0604 die Endpunkterheber nicht verblindet, jedoch erhielten in dieser Phase vor der Randomisierung alle Patienten Verum (Wirkstoffkombination THC:CBD 1:1). Eine Verzerrung bei der Endpunkterhebung zu diesem Zeitpunkt dieser Studie ist deshalb ebenfalls entsprechend unwahrscheinlich. Der Endpunkt wurde, wie geplant, für die jeweilige ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt, wie geplant, statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in allen beiden Studien das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

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

Tabelle 4-43: Ergebnisse für Aktivitäten des täglichen Lebens (Barthel Index) aus RCT mit dem zu bewertenden Arzneimittel

Studie	Aktivitäten des täglichen Lebens (Barthel Index Gesamt Score)								
	Baseline			Studienende			Veränderung		
	N	Mean (SD)	Median Min- Max	N	Mean (SD)	Median Min-Max	Mean (SD)	Median Min-Max	Gruppen- differenz (*) (95%-CI) p-Wert
GWCL0403 ¹⁾									
Sativex®	166	70,0 (26,33)	80 (0-100)	162	70,9 (27,32)	80 (0-100)	-0,1 (10,26)	0,0 (-65-20)	-0,15 (-1,95; 1,64) p=0,8665
Placebo	169	72,0 (23,29)	80 (0-100)	165	72,8 (23,19)	80 (0-100)	0,5 (8,05)	0,0 (-20-35)	
GWSP0604 ²⁾									
Sativex®	122	73,8 (24,87)	85 (5-100)	121	74,0 (24,72)	80 (10-100)	0 (7,71)	0 (-25-30)	1,84 (0,02; 3,66) p=0,0479
Placebo	116	78,1 (23,34)	85 (0-100)	117	76,4 (24,04)	85 (0-100)	-1,9 (6,34)	0 (-25-10)	
(*) Adjustiertes ANCOVA Modell									
1. Verbesserung 14 Wochen nach Behandlungsbeginn									
2. Verbesserung 12 Wochen nach Fortführung versus Absetzen nach Ansprechen (mindestens 20 % Verbesserung der NRS) auf 4 Wochen Run-in Behandlung mit Sativex®									

Die Aktivitäten des täglichen Lebens wurden mittels Barthel Index in zwei RCT erhoben. In einer Studie (GWCL0403) zeigte sich im direkten Vergleich wenig Unterschied zwischen der Therapiefortsetzung ohne THC:CBD 1:1 (Placebogruppe) und der Therapiefortsetzung mit THC:CBD 1:1, die Gruppendifferenz für die Gesamtpunktzahl des Barthel Index betrug in dieser Studie -0,15 (95 %-CI: -1,95; 1,64). In der zweiten Studie GWSP0604, die dem Zulassungsstatus von Sativex® am besten entspricht, also den Patienten, die auf eine Therapie mit Sativex® in klinisch relevanter Form ansprechen, ergab sich eine auf dem 5 % Niveau statistisch signifikante Differenz von 1,84 (95 %-CI: 0,02, 3,66) zugunsten der Therapiefortsetzung mit THC:CBD 1:1 und entspricht nahezu der klinisch relevanten Veränderung des Barthel Index von 1,85 Punkten (19).

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.

Eine Meta-Analyse wurde nicht durchgeführt, da in den Studien unterschiedliche Patientenpopulationen in einem unterschiedlichen Behandlungsmuster betrachtet wurden. In Studie GWCL0403 wurden naive Patienten im Follow-up mit Sativex[®] oder Placebo behandelt, in der Responderstudie GWSP0604 wurde das Fortführen von Sativex[®] versus Absetzen (Placebogruppe) bei Patienten untersucht, die zuvor bereits auf eine 4-wöchige Sativex[®] Behandlung positiv angesprochen hatten.

4.3.1.3.1.11 Gehfähigkeit (10 m-Gehstreckentest) – 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.

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

Tabelle 4-44: Operationalisierung von Gehfähigkeit (10 m-Gehstreckentest)

Studie	Operationalisierung
GWCL0403	Die Gehfähigkeit wurde als Zeit für das Zurücklegen einer geraden Strecke von 10 m durch den Investigator erfasst.
GWSP0604	Die Gehfähigkeit wurde als Zeit für das Zurücklegen einer geraden Strecke von 10 m durch den Investigator erfasst.

Für die Studie GWMS0001 wurde die Gehfähigkeit mit Hilfe des 10 m-Gehstreckentests untersucht. Allerdings erfolgte dabei keine Auswertung für die 37 Patienten mit primärer Spastik. Aus diesem Grund wurde diese Studie für den Endpunkt Gehfähigkeit nicht berücksichtigt.

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 4-45: Bewertung des Verzerrungspotenzials für Gehfähigkeit (10 m-Gehstreckentest) 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
GWCL0403	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWSP0604	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Aufgrund des niedrigen Verzerrungspotentials auf Studienebene aller relevanten RCT, die den Endpunkt Gehfähigkeit (10 m-Gehstreckentest) untersucht haben, kann von einem niedrigen Verzerrungspotential auch auf Ebene des Endpunkts selbst ausgegangen werden. In den beiden Studien waren die Endpunkterheber (Study Investigator) am Zeitpunkt der Endpunkterhebung verblindet. Zwar waren in der ersten Phase in Studie GWSP0604 die Endpunkterheber nicht verblindet, jedoch erhielten in dieser Phase vor der Randomisierung alle Patienten Verum (Wirkstoffkombination THC:CBD 1:1). Eine Verzerrung bei der Endpunkterhebung zu diesem Zeitpunkt dieser Studie ist deshalb ebenfalls entsprechend unwahrscheinlich. In allen Studien wurde der Endpunkt, wie geplant, für die jeweilige ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt, wie geplant, statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in beiden Studien das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

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

Tabelle 4-46: Ergebnisse für Gehfähigkeit (10 m-Gehstreckentest) aus RCT mit dem zu bewertenden Arzneimittel

Studie	10 m-Gehstreckentest (Sekunden)								
	Baseline			Studienende			Veränderung		
	N	Mean (SD)	Median Min-Max	N	Mean (SD)	Median Min-Max	Mean (SD)	Median Min-Max	Gruppen- differenz (*) (95%-CI) p-Wert
GWCL 0403 ¹⁾									
Sativex®	166	34,8 (60,55)	17 (7-502)	115	31,4 (59,15)	14 (6-490)	-2,1 (17,37)	-1,0 (-114-55)	0,0 (-2,0;-1,0) p=0,6240
Placebo	169	36,7 (70,77)	16,5 (6-607)	120	46,2 (119,91)	14,5 (7-857)	9,3 (63,56)	-0,5 (-37-514)	
GWSP 0604 ²⁾									
Sativex®	87	24,5 (25,51)	15 (5-135)	84	20,0 (16,84)	14 (5-93)	-2,3 (8,38)	0 (-35-13)	-3,34 (-6,95;-0,26) p=0,0687
Placebo	86	25,3 (36,20)	15 (7-306)	83	25,4 (37,51)	15 (7-292)	2,0 (15,16)	0 (-14-119)	
(*) Adjustiertes ANCOVA Modell									
1. Veränderung 14 Wochen nach Behandlungsbeginn									
2. Veränderung 12 Wochen nach Fortführung versus Absetzen nach Ansprechen (mindestens 20 % Verbesserung der NRS) auf 4 Wochen Run-in Behandlung mit Sativex®									

Der 10 m-Gehstreckentest wurde in zwei RCT durchgeführt. In einer Studie (GWCL0403) zeigte sich im direkten Vergleich kein Unterschied zwischen der Therapiefortsetzung ohne THC:CBD 1:1 (Placebogruppe) und der Therapiefortsetzung mit THC:CBD 1:1, die Gruppendifferenz für die Gesamtpunktzahl der 10 m-Gehstrecke betrug in dieser Studie 0,0 (95 %-CI: -2,0, 1,0). In der anderen Studie (GWSP0604) ergab sich eine Differenz von -3,34 (95 %-CI: -6,95, 0,26), die statistisch nicht signifikant ist. Die Änderung der Geschwindigkeit beträgt in der Studie GWSP0604 3,34 s auf 10 m bei der Add-On Behandlung mit THC:CBD 1:1 gegenüber Placebo. Die Geschwindigkeitsänderung in der Studie GWSP0604 im Studienarm THC:CBD 1:1 beträgt 0,092 m/s (10 m/20 s am Studienende minus 10 m/24,5 s Baseline) und liegt somit über der klinisch relevanten Verbesserung der Gehfähigkeit von 0,05 m/s.(22).

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.

Eine Meta-Analyse wurde nicht durchgeführt, da in den Studien unterschiedliche Patientenpopulationen in einem unterschiedlichen Behandlungsmuster betrachtet wurden. In der Studie GWCL0403 wurden naive Patienten im Follow-up mit Sativex[®] oder Placebo behandelt, in der Responderstudie GWSP0604 wurde das Fortführen von Sativex[®] versus Absetzen (Placebogruppe) bei Patienten untersucht, die zuvor bereits auf eine 4-wöchige Sativex[®] Behandlung positiv angesprochen hatten.

4.3.1.3.1.12 Nebenwirkungen und schwerwiegende Nebenwirkungen (mit möglichem Zusammenhang zur Studienmedikation) – 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.

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

Tabelle 4-47: Operationalisierung von Nebenwirkungen

Studie	Operationalisierung
GWMS0106	Nebenwirkungen und schwerwiegende Nebenwirkungen wurden über den gesamten Studienzeitraum GCP-konform dokumentiert bis einschließlich der Follow-up Periode nach Studienende nach Anzahl der Patienten mit Ereignis.
GWCL0403	Nebenwirkungen und schwerwiegende Nebenwirkungen wurden über den gesamten Studienzeitraum GCP-konform dokumentiert bis einschließlich der Follow-up Periode nach Studienende nach Anzahl der Patienten mit Ereignis.
GWSP0604	Nebenwirkungen und schwerwiegende Nebenwirkungen wurden über den gesamten Studienzeitraum GCP-konform dokumentiert bis einschließlich der Follow-up Periode nach Studienende separat für Phase A und B nach Anzahl der Patienten mit Ereignis.
ARA2009	Nebenwirkungen wurden über den gesamten Studienzeitraum GCP-konform dokumentiert bis einschließlich einer 2-wöchigen Follow-up Periode nach Studienende nach Anzahl der Ereignisse.

Für die Studie GWMS0001 wurden die Nebenwirkungen untersucht. Allerdings erfolgten dabei bei der Auswertung keine eigenständigen Angaben für die 37 Patienten mit primärer Spastik. Aus diesem Grund wurde diese Studie für den Endpunkt schwerwiegende Nebenwirkungen nicht berücksichtigt.

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 4-48: Bewertung des Verzerrungspotenzials für Nebenwirkungen mit möglichem Zusammenhang zur Studienmedikation 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
GWMS0106	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWCL0403	Niedrig	Ja	Ja	Nein	Nein	Niedrig
GWSP0604	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ARA2009	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Aufgrund des niedrigen Verzerrungspotentials auf Studienebene aller relevanten RCT, die den Endpunkt Nebenwirkungen untersucht haben, kann von einem niedrigen Verzerrungspotential auch auf Ebene des Endpunkts selbst ausgegangen werden. In allen vier Studien waren die Endpunkterheber (Study Investigator) in der doppelblinden randomisierten Gruppenphase verblindet. Der Endpunkt wurde, wie geplant, für die jeweilige Safety Population (alle Patienten, die mindestens eine Studienmedikation bekommen hatten) ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt wie geplant statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in allen Studien das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

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

Tabelle 4-49: Ergebnisse für Nebenwirkungen aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anzahl Patienten mit Nebenwirkungen				Anzahl der Nebenwirkungen	
	THC:CBD 1:1		Placebo		THC:CBD 1:1	Placebo
	N	n (%)	N	n (%)		
GWMS0106	124	102 (82)	65	46 (71)	169	49
GWCL0403	167	156 (93)	170	132 (78)	357	216
GWSP0604	124	66 (53)	117	57 (49)	112	88
ARA 2009	17	ND	17	ND	41	11

Die vorliegenden Ergebnisse zeigen, dass in den Studie GWMS0106, GWCL0403 und GWSP0604 in beiden Behandlungsarmen Nebenwirkungen auftraten, wobei mehr Patienten Nebenwirkungen aufwiesen, denen die Wirkstoffkombination THC:CBD 1:1 als Add-on Therapie verabreicht wurde. Dies lässt aber keine eindeutige Aussage zur Sicherheit des Arzneimittels zu. Deshalb wurde eine weitere Analyse zur Häufigkeit des Auftretens von schwerwiegenden Nebenwirkungen mit möglichem Zusammenhang zur Studienmedikation durchgeführt.

Tabelle 4-50: Ergebnisse für Schwerwiegende Nebenwirkungen mit möglichem Zusammenhang zur Studienmedikation aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anzahl Patienten mit Schwerwiegenden Nebenwirkungen				Anzahl Patienten mit Schwerwiegenden Nebenwirkungen mit möglichem Zusammenhang zur Studienmedikation	
	THC:CBD 1:1		Placebo		THC:CBD 1:1	Placebo
	N	n (%)	N	n (%)	n (%)	n (%)
GWMS0106	124	4 (3,2)	65	3 (4,6)	1 (0,8)	0 (0,0)
GWCL0403	167	4 (2,0)	170	1 (1,0)	4 (2,0)	1 (1,0)
GWSP0604	124	8 (6,5)	117	3 (2,6)	1 (0,8)	0 (0,0)
ARA 2009	17	0 (0)	17	0 (0)	0 (0,0)	0 (0,0)

Die vorliegenden Analysenergebnisse aus allen vier Studien zeigen eine geringe Häufigkeit für das Auftreten von schwerwiegenden Nebenwirkungen mit möglichem Zusammenhang zur Studienmedikation. In einer Studie (ARA2009) wurde unter THC:CBD 1:1 keine und in zwei weiteren Studien (GWMS0106, GWSP0604) jeweils nur von einem Patient mit schwerwiegenden Nebenwirkungen mit möglichem Zusammenhang zur Studienmedikation

berichtet (Placebo 0 (0 %)). In der Studie GWCL0403 wurden unter THC:CBD 1:1 von vier (2,0 %) Patienten (Placebo: 1(1,0 %) mit schwerwiegenden Nebenwirkungen berichtet. Unter der Wirkstoffkombination THC:CBD 1:1 wurden folgende schwerwiegende Nebenwirkungen berichtet: Reizbarkeit (N=1), Harnwegsinfekt (N=1), Muskelspasmen (N=1), Aggression (N=1), Agitation (N=1), Verwirrtheit (N=1), Depression (N=1), Drogenabhängigkeit (N=1), Insomnie (N=1), Paranoia (N=1), Suizidgedanken (N=1). Unter Placebo wurde Harnverhalt (N=1) berichtet.

Auch die intrathekale Gabe von Baclofen hat (wie Sativex[®]) eine Zulassung bei der Behandlung „bei schwerer, chronischer Spastizität bei Multipler Sklerose“ (40). Und wie bei Sativex[®] kommt diese Therapie nur dann in Betracht, wenn andere Behandlungsmöglichkeiten ausgeschöpft wurden („die mit einer medikamentösen Standardtherapie nicht erfolgreich behandelt werden kann“ (40)). Auf Grund der Schwere und Dramatik der Nebenwirkungen (41) sowie der hohen Invasivität dieser Therapieform (Implantation einer Pumpe mit Verbindung zum Liquorraum) wird diese Therapieoption in aller Regel nur bei sehr schweren Verläufen angewandt.

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.

Eine Meta-Analyse wurde für diesen Endpunkt nicht durchgeführt, da dieser Endpunkt in allen Studien explorativ erhoben wurde.

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.

Bezüglich der patientenrelevanten Endpunkte wurden keine Subgruppen analysiert, da die Ergebnisse in den einzelnen Studien nicht separat nach Alter, Geschlecht oder Schwere der Krankheit dargestellt wurden.

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.

Der Zusatznutzen einer Zusatzbehandlung mit Sativex[®] von Patienten mit mittelschwerer bis schwerer Spastik aufgrund von Multipler Sklerose (MS), die nicht angemessen auf eine andere anti-spastische Arzneimitteltherapie angesprochen haben und die eine klinisch erhebliche Verbesserung von mit der Spastik verbundenen Symptomen während eines Anfangstherapieversuchs aufzeigen, ergibt sich aus dem direkten Vergleich mit der Fortführung der anti-spastischen Vortherapie plus Placebo.

Unabhängig von den nach dem derzeitigen Stand der Medizin existierenden Behandlungsmöglichkeiten der Spastik mit Baclofen, Tizanidin oder verschiedenen Benzodiazepinen in Mono- oder Kombinationstherapie zeigen die angeführten RCT einen beträchtlichen Zusatznutzen bei einer Zusatzbehandlung mit Sativex[®].

Der beträchtliche Zusatznutzen der Zusatzbehandlung mit Sativex[®] zeigt sich in einer höheren Rate von Patienten mit klinisch relevanter Verbesserung der Spastik, einer Reduktion der durchschnittlichen Anzahl an Spastiken pro Tag, einer Verbesserung der Schlafqualität, einer besseren globalen Beurteilung der Veränderung durch die Patienten, die betreuenden Personen und die behandelnden Ärzte, einen besseren Gesundheitszustand, einer Verbesserung der Aktivitäten des täglichen Lebens und der Gehfähigkeit. Die Häufigkeit von schwerwiegenden Nebenwirkungen mit möglichem Zusammenhang zur Behandlung mit THC:CBD 1:1 war gering.

Im direkten Vergleich zur Therapiefortsetzung ohne die Wirkstoffkombination THC:CBD 1:1 zeigt die Zusatztherapie unter THC:CBD 1:1 in allen beurteilten RCT eine deutlichere Verbesserung der Spastik auf der 11-stufigen NRS bzw. VAS. In drei der vier Studien (GWMS0001, GWMS0106 und GWSP0604) war die Mittelwertdifferenz im Vergleich zur Therapiefortsetzung ohne THC:CBD 1:1 statistisch signifikant zugunsten der Zusatztherapie mit THC:CBD 1:1 (Tabelle 4-51). Die Meta-Analyse dieses Endpunkts aus den drei vergleichbaren Studien GWMS0001, GWMS0106 und GWCL0403 zeigte einen statistischen Vorteil der Wirkstoffkombination THC:CBD 1:1 gegenüber Placebo (Abbildung 4).

Tabelle 4-51: Zusammenfassung der Ergebnisse zur Verbesserung des Grades der Spastik auf der NRS

Studie	Mittelwertdifferenz [95 % CI]	p-Wert	Meta-Analyse OR [95%-CI]
GWMS0001	-1,84 [3,181; -0,501]	0,009	-0,24 [-0,42; -0,05] p=0,01 zugunsten von THC:CBD 1:1
GWMS0106	-0,517 [-1,029; -0,004]	0,048	
GWCL0403	-0,17 [- 0,53; 0,19]	0,3592	
GWSP0604	-0,84 [- 1,29; -0,4]	0,0002	

Der Anteil der Patienten mit klinisch relevanter Verbesserung ist unter Zusatzbehandlung mit THC:CBD 1:1 höher (Tabelle 4-52).

Tabelle 4-52: Zusammenfassung der Ergebnisse zur klinisch relevanten Verbesserung der Spastik auf der NRS

Studie	Anteil der Patienten mit ≥ 30 % Verbesserung der Schwere der Spastik		Anteil der Patienten mit ≥ 50 % Verbesserung der Schwere der Spastik	
	OR [95 %CI]	p-Wert	OR [95 %CI]	p-Wert
GWMS0001	8,57 [1,82; 40,42]	0,007	3,04 [0,51; 18,11]	0,22
GWMS0106	2,38 [1,19; 4,78]	0,01	2,05 [0,78; 5,37]	0,14
GWCL0403	1,34 [0,83; 2,17]	0,23	1,21 [0,62; 2,37]	0,569
GWSP0604	2,73 [1,59; 4,69]	0,0003	1,65 [0,98; 2,78]	0,061

Eine Meta-Analyse dieses Endpunkts aus den Ergebnissen der drei RCT (GWMS001, GWMS0106 und GWCL0403) für eine Verbesserung von mindestens 30 % auf der NRS ergab eine Odds ratio von 1,69 zugunsten von THC:CBD 1:1 ($p = 0,003$). In der RCT mit Patienten mit der zugelassenen Indikationsstellung für Sativex[®], d.h. Patienten, die auf

Sativex[®] ansprechen (GWSP0604), betrug die Odds ratio bei einer 30 %igen Verbesserung auf der NRS zur Vergleichsgruppe 2,73 (95 %-CI: 1,59; 4,69) zugunsten von Sativex[®] und war statistisch signifikant ($p = 0,0003$). Eine Sensitivitätsanalyse ergab für die Verbesserung von mindestens 50 % auf der NRS für die Studien GWMS0001, GWMS0106 und GWCL0403 einen numerischen, nicht signifikanten ($p=0,1$) Vorteil der Add-on Therapie mit THC:CBD 1:1 gegenüber Placebo (OR=1,55).

Die durchschnittliche Anzahl der Spastiken war in der Responderstudie GWSP0604, d.h. bei Patienten, die genau der zugelassenen Indikation entsprechen, um durchschnittlich -2,53 Spastiken pro Tag (95 %-CI: -4,27, -0,79) signifikant geringer unter Sativex[®] gegenüber der Vergleichstherapie ($p = 0,004$) (Tabelle 4-53).

Tabelle 4-53: Zusammenfassung der Ergebnisse zur Anzahl der Spastiken

Studie	Anzahl der Spastiken pro Tag		Durchschnittliche Anzahl Spastiken pro Tag	
	Mittelwertdifferenz [95%-CI]	p-Wert	Mittelwertdifferenz [95%-CI]	p-Wert
GWMS0106	-0,17 [-0,39; 0,06]	0,141	_____	_____
GWMS0604	_____	_____	-2,53 [-4,27;-0,79]	0,004

Die Schlafqualität zeigte sich ebenfalls stärker verbessert unter der Zusatzbehandlung mit der Wirkstoffkombination THC:CBD 1:1. In der zugelassenen Indikationsstellung für Sativex[®], d.h. bei Patienten, die auf Sativex[®] ansprechen (Studie GWSP0604), betrug die Differenz zur Vergleichsgruppe -0,88 (95 %-CI: (-1,25; -0,51) zugunsten von THC:CBD 1:1 und war statistisch signifikant ($p < 0,001$) (Tabelle 4-25).

Der Gesundheitszustand, gemessen mittels EQ-5D, zeigte sich in zwei Studien zwischen den Vergleichsgruppen nur wenig verändert. Insgesamt war die Zunahme unter der Zusatzbehandlung mit THC:CBD 1:1 leicht verbessert (Studie GWCL0403), bzw. die Abnahme weniger stark ausgeprägt im Vergleich zum Absetzen der Studienmedikation (Responderstudie GWSP0604). Die Differenz zwischen den Behandlungsgruppen für die Veränderung des 100 mm VAS-Werts betrug 1,42 (95 %-CI: -3,11; -5,94) und 1,24 (95 %-CI: -3,01; -5,5) zugunsten von THC:CBD 1:1 (Tabelle 4-28).

Die Lebensqualität wurde in drei Studien (GWCL0403, GWSP0604 und ARA2009) untersucht, dabei wurden aber drei verschiedene Messmethoden angewendet. Aus den vorliegenden Ergebnissen lassen sich keine eindeutigen Aussagen treffen, ob und in welchen Bereichen sich eine Verbesserung der Lebensqualität bei der Add-on Therapie einstellt (Tabelle 4-31).

Die Globalbeurteilung der klinischen Veränderung durch verschiedene Personen zeigte für Patienten in der zugelassenen Indikationsstellung für Sativex[®], (d.h. bei Patienten, die auf Sativex[®] ansprechen), eine konsistente, statistisch signifikante Verbesserung unter der Zusatzbehandlung mit THC:CBD 1:1. Für die Globalbeurteilung der Veränderung, bewertet durch den Patienten selbst, ergab sich in der Studie GWSP0604 eine Odds ratio für eine kategorielle Verbesserung der Globalbeurteilung von 1,703 (95 %-CI: 1,075; 2,698) zugunsten von THC:CBD 1:1 (Tabelle 4-34). Für die Globalbeurteilung der Veränderung der funktionalen Möglichkeiten des Patienten durch die betreuende Person ergab sich in der Studie GWSP0604 eine Odds ratio für eine kategorielle Verbesserung der Globalbeurteilung von 2,4 (95 %-CI: 1,29; 4,44) zugunsten von THC:CBD 1:1 (Tabelle 4-37). Für die Globalbeurteilung der Veränderung der Transferierbarkeit des Patienten durch die betreuende Person ergab sich in der Studie GWSP0604 eine jeweils statistisch auf dem 5 % Niveau nicht signifikante Odds ratio für eine kategorielle Verbesserung der Globalbeurteilung von 1,792 (95 %-CI: 0,973, 3,301) zugunsten von THC:CBD 1:1. Für die Globalbeurteilung der Veränderung des Patienten durch den behandelnden Arzt ergab sich in der Studie GWSP0604 eine statistisch signifikante Odds ratio für eine kategorielle Verbesserung der Globalbeurteilung von 1,958 (95 %-CI: 1,232, 3,112) zugunsten von THC:CBD 1:1 (Tabelle 4-40).

Der Barthel Index als Maß für die Aktivitäten des täglichen Lebens war unter der Zusatzbehandlung mit THC:CBD 1:1 vergleichsweise besser (Tabelle 4-43). In der gemischten Populationen von Patienten (Studie GWCL0403) war der Unterschied gering, die Differenz zur Vergleichsgruppe betrug -0,15 (95 %-CI: -1,95, 1,64). Unter Patienten mit der zugelassenen Indikationsstellung für Sativex[®], d.h. Patienten, die auf Sativex[®] ansprechen (Studie GWSP0604), betrug die Differenz zur Vergleichsgruppe 1,84 (95 %-CI: 0,02; 3,66) zugunsten von Sativex[®] und war statistisch signifikant ($p = 0,0479$).

Die Mobilität der Patienten war im 10 m-Gehstreckentest unter Zusatzbehandlung mit der Wirkstoffkombination THC:CBD 1:1 vergleichsweise besser (Tabelle 4-46). In der gemischten Populationen von Patienten (Studie GWCL0403) war kein Unterschied festzustellen. Unter Patienten mit der zugelassenen Indikationsstellung für Sativex[®], d.h. Patienten, die auf Sativex[®] ansprechen (Studie GWSP0604), betrug die Differenz zur Vergleichsgruppe -3,34 s (95 %-CI: -6,95; 0,26) zugunsten von THC:CBD 1:1. Die Differenz war auf dem 5 % Niveau zwar statistisch nicht signifikant, jedoch mit 0,092 m/s für den Studienarm THC:CBD 1:1 deutlich über der klinischen Relevanz von 0,05 m/s.

Die Auswertung der Nebenwirkungen zeigte in den ausgewerteten Studien für beide Behandlungsarme eine höhere Anzahl von Nebenwirkungen in der THC:CBD 1:1-Gruppe (Tabelle 4-49). Die deskriptiv ausgewertete Häufigkeit des Auftretens schwerwiegender Nebenwirkungen zeigt eine geringe Anzahl von Patienten mit möglichem Zusammenhang zur Studienmedikation in allen Studien. In einer Studie (ARA2009) wurde unter Sativex[®] keine und in zwei weiteren Studien (GWMS0106, GWSP0604) jeweils nur von einem Patient mit schwerwiegenden Nebenwirkungen mit möglichem Zusammenhang zur Studienmedikation

berichtet (Placebo 0 (0 %)). In der Studie GWCL0403 wurden unter Sativex[®] von vier (2,0 %) Patienten (Placebo: 1 (1,0 %) mit schwerwiegenden Nebenwirkungen berichtet.

Alle Autoren der Studienberichte kamen zu dem Ergebnis, dass das Oromukosalspray mit der Wirkstoffkombination THC:CBD 1:1 nur wenig schwerwiegende Nebenwirkungen aufzeigt und außerdem eine gute Verträglichkeit aufweist.

In Tabelle 4-54 sind noch einmal alle Endpunkte aufgeführt und auf einer selbst definierten Skala für jede RCT bewertet worden. Aus dieser Tabelle kann man schließen, dass die Add-on Therapie in allen funktionellen patientenrelevanten Endpunkten (Spastik, Aktivitäten des täglichen Lebens und der klinischen Globalbeurteilung durch den Arzt) einen Vorteil gegenüber der Vergleichstherapie aufweist.

Tabelle 4-54: Zusammenfassende Einschätzung aller Endpunkte

Studie/ Endpunkt	GWMS 0001	GWMS 0106	GWCL 0403	GWSP 0604	ARA 2009
Schweregrad der Spastik	++	++	+	++	
Anzahl Patienten mit Verbesserung der Schwere der Spastik ($\geq 30\%$)	++	++	+	++	
Anzahl Spastiken pro Tag		+		++	
Schlafqualität			+	++	
Gesundheitszustand			+	+	
Lebensqualität			±	±	±
Aktivitäten des täglichen Lebens			+	++	
Gehfähigkeit			+	+	
Klinische Globalbeurteilung		+	+	++	
		(nur SGIC)	(nur CGIC)		
Nebenwirkungen		±	±	±	±
++ statistisch signifikanter Vorteil von THC:CBD 1:1 gegenüber Placebo + numerischer Vorteil von THC:CBD 1:1 gegenüber Placebo ±.....kein Vorteil von THC:CBD 1:1 gegenüber Placebo leer Endpunkt in dieser Studie nicht erhoben/ausgewertet					

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*

ND – Es wurde kein indirekter Vergleich durchgeführt.

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.

ND – Es wurde kein indirekter Vergleich durchgeführt.

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 4-55: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

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

4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

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

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

Tabelle 4-56: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

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

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

ND – Es wurde kein indirekter Vergleich durchgeführt.

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

Tabelle 4-57: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-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 4-58: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> 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
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

ND – Es wurde kein indirekter Vergleich durchgeführt.

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 4-59: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

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

ND – Es wurde kein indirekter Vergleich durchgeführt.

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.*

ND – Es wurde kein indirekter Vergleich durchgeführt.

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.

ND – Es wurde kein indirekter Vergleich durchgeführt.

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

- 1. Studien des pharmazeutischen Unternehmers*
- 2. Studien aus der bibliografischen Literaturrecherche*
- 3. Studien aus der Suche in Studienregistern*
- 4. Resultierender Studienpool aus den einzelnen Suchschritten*

ND – Es wurden keine nicht randomisierten vergleichenden Studien in diesem Dossier betrachtet.

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 4-60: 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>

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 4-61: 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 4-62: 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

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

ND – Es wurden keine nicht randomisierten vergleichenden Studien in diesem Dossier betrachtet.

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.

ND – Es wurden keine nicht randomisierten vergleichenden Studien in diesem Dossier betrachtet.

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

- 1. Studien des pharmazeutischen Unternehmers*
- 2. Studien aus der bibliografischen Literaturrecherche*
- 3. Studien aus der Suche in Studienregistern*
- 4. Resultierender Studienpool aus den einzelnen Suchschritten*

Tabelle 4-63: Liste der Studien des pharmazeutischen Unternehmers – weitere Untersuchungen mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studien-dauer	Therapiearme
GWCADD01 A safety and tolerability study of medicinal cannabis extracts for chronic refractory spasticity and neurogenic pain	Nein	abgeschlossen	11 Wochen 08/2001 - 05/2002	GW-1000/1-02: THC/CBD, extract of Cannabis sativa L. GW-2000-02 THC extract of Cannabis sativa L. GW-3000-02 CBD extract of Cannabis sativa L. Matching placebo:GW-4001-01
GWN 10005 An “N of 1” placebo controlled study of cannabis based medicine extracts in treatment of multiple sclerosis, spinal cord injury, peripheral nerve injury or rheumatoid arthritis.	Nein	abgebrochen	12 Wochen 10/2001– 11/2001	GW-1000/1-02: Δ9-tetrahydro- cannabinol THC/CBD extract of Cannabis sativa L. GW-2000-02 THC extract of Cannabis sativa L. GW-3000-02 CBD extract of Cannabis sativa L. Matching placebo: GW-4001-01
GWN19902 A preliminary double-blind, randomised crossover, placebo-controlled investigation of the therapeutic profile of three cannabis-based medicine extracts in patients with multiple sclerosis, spinal cord conditions, peripheral nerve injury or central nervous system damage associated with	Nein	abgeschlossen	10 Wochen 07/2000 - 11/2001	GW-1000/1/2 THC/CBD extract of Cannabis sativa L. GW-2000 THC extract of Cannabis sativa L. GW-3000 CBD extract of Cannabis sativa L.

Studienbezeichnung	Zulassungsstudie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studien-dauer	Therapiearme
vascular, traumatic, infective, genetic, neoplastic and metabolic disease; any of which has resulted in pain and/or spasticity ⁷				Placebo
GWN19904 A preliminary placebo-controlled investigation of the therapeutic profile of three cannabis based medicine extracts in patients with multiple sclerosis and other neurological conditions resulting in pain and/or spasticity, or arthritic conditions	Nein	abgeschlossen	8 Wochen 09/2000-03/2002	GW-1000/1 THC/CBD extract of Cannabis sativa L. GW-2000/1 THC extract of Cannabis sativa L. GW-3001/2 CBD extract of Cannabis sativa L. GW-4000/1: Placebo
GWMS0001 A double blind, randomised, parallel group, placebo-controlled trial of a combination of delta-9-tetrahydrocannabinol (THC) and cannabidiol (CBD) in patients with multiple sclerosis, followed by an open label assessment and study extension.	Nein	abgeschlossen	6 Wochen 05/2001 - 07/2002	GW-1000-02 THC/CBD extract of Cannabis sativa L Placebo
GWMS0106 A Double Blind, Randomised, Parallel Group Study to Assess the Efficacy, Safety and Tolerability of Cannabis Based Medicine 1:1 THC/CBD Compared With Placebo for the Treatment of Spasticity in Patients With Multiple Sclerosis.	Ja	abgeschlossen	8 Wochen (2 Wochen Baseline + 6 Wochen treatment) 06/2002 – 03/2004	THC/CBD (Sativex®) Placebo
GWSP0604 A two-phase, phase 3 study of the safety and efficacy of Sativex®, in the symptomatic relief of spasticity in subjects with spasticity due to multiple sclerosis: PHASE A – Single-blind response assessment; PHASE B – Double-blind randomized, placebo controlled, parallel group study	Ja	abgeschlossen	19 Wochen 01/2008-01/2009	THC/CBD (Sativex®) Placebo
GWSP0702 A Placebo Controlled, Parallel Group, Randomised Withdrawal	Ja	abgeschlossen	5 Wochen 11/2007-	THC/CBD (Sativex®)

Studienbezeichnung	Zulassungs- studie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studien- dauer	Therapiearme
Study of Subjects With Symptoms of Spasticity Due to Multiple Sclerosis Who Are Receiving Long-term GW-1000-02 (Sativex®).			01/2009	Placebo
GWCL0403 A double blind, randomised, placebo controlled, parallel group study of Sativex®, in subjects with symptoms of spasticity due to multiple sclerosis	Ja	abgeschlossen	14 Wochen 03/2005- 12/2005	THC/CBD (Sativex®) Placebo

Stand der Information: 1. November 2011.

Tabelle 4-64: Studien des pharmazeutischen Unternehmers, die nicht für weitere Untersuchungen mit dem zu bewertenden Arzneimittel in Frage kommen

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
<p>GWCADD01</p> <p>A safety and tolerability study of medicinal cannabis extracts for chronic refractory spasticity and neurogenic pain</p>	<p>Das Einschlusskriterium E3 ist nicht erfüllt. Nur 2 von 8 Patienten sind MS-Patienten.</p>
<p>GWN 10005</p> <p>An “N of 1” placebo controlled study of cannabis based medicine extracts in treatment of multiple sclerosis, spinal cord injury, peripheral nerve injury or rheumatoid arthritis.</p>	<p>Das Einschlusskriterium E3 ist nicht erfüllt. Es handelte sich nicht um einen MS-Patienten.</p>
<p>GWN19902</p> <p>A preliminary double-blind, randomised crossover, placebo-controlled investigation of the therapeutic profile of three cannabis-based medicine extracts in patients with multiple sclerosis, spinal cord conditions, peripheral nerve injury or central nervous system damage associated with vascular, traumatic, infective, genetic, neoplastic and metabolic disease; any of which has resulted in pain and/or spasticity⁷</p>	<p>Das Einschlusskriterium E4 ist nicht erfüllt. Zu den 17 MS-Patienten von insgesamt 25 Patienten liegt keine Unterscheidung zur Stärke der Spastizität in leicht, mittelschwer und schwer vor. Außerdem erfolgte keine separate Auswertung der Ergebnisse für diese Patienten.</p>
<p>GWN19904</p> <p>A preliminary placebo-controlled investigation of the therapeutic profile of three cannabis based medicine extracts in patients with multiple sclerosis and other neurological conditions resulting in pain and/or spasticity, or arthritic conditions</p>	<p>Das Einschlusskriterium E4 ist nicht erfüllt. Zu den 14 MS-Patienten von insgesamt 27 Patienten liegt keine Unterscheidung zur Stärke der Spastizität in leicht, mittelschwer und schwer vor.</p>
<p>GWMS0001</p> <p>A double blind, randomised, parallel group, placebo-controlled trial of a combination of delta-9-tetrahydrocannabinol (THC) and cannabidiol (CBD) in patients with multiple sclerosis, followed by an open label assessment and study extension.</p>	<p>Die Studie wurde für die Nutzenbewertung herangezogen.</p>
<p>GWMS0106</p> <p>A Double Blind, Randomised, Parallel Group Study to Assess the Efficacy, Safety and Tolerability of Cannabis Based Medicine 1:1 THC/CBD Compared With Placebo for the Treatment of Spasticity in Patients With Multiple Sclerosis.</p>	<p>Die Studie wurde für die Nutzenbewertung herangezogen.</p>
<p>GWSP0604</p> <p>A two-phase, phase 3 study of the safety and efficacy of Sativex[®], in the symptomatic relief of spasticity in subjects with spasticity due to multiple sclerosis: PHASE A – Single-blind response assessment; PHASE B – Double-blind randomized, placebo controlled, parallel group study</p>	<p>Die Studie wurde für die Nutzenbewertung herangezogen.</p>
<p>GWCL0403</p> <p>A double blind, randomised, placebo controlled, parallel group study of Sativex[®], in subjects with symptoms of spasticity due to multiple sclerosis</p>	<p>Die Studie wurde für die Nutzenbewertung herangezogen.</p>

4.3.2.3.2 Studien aus der bibliografischen Literaturrecherche

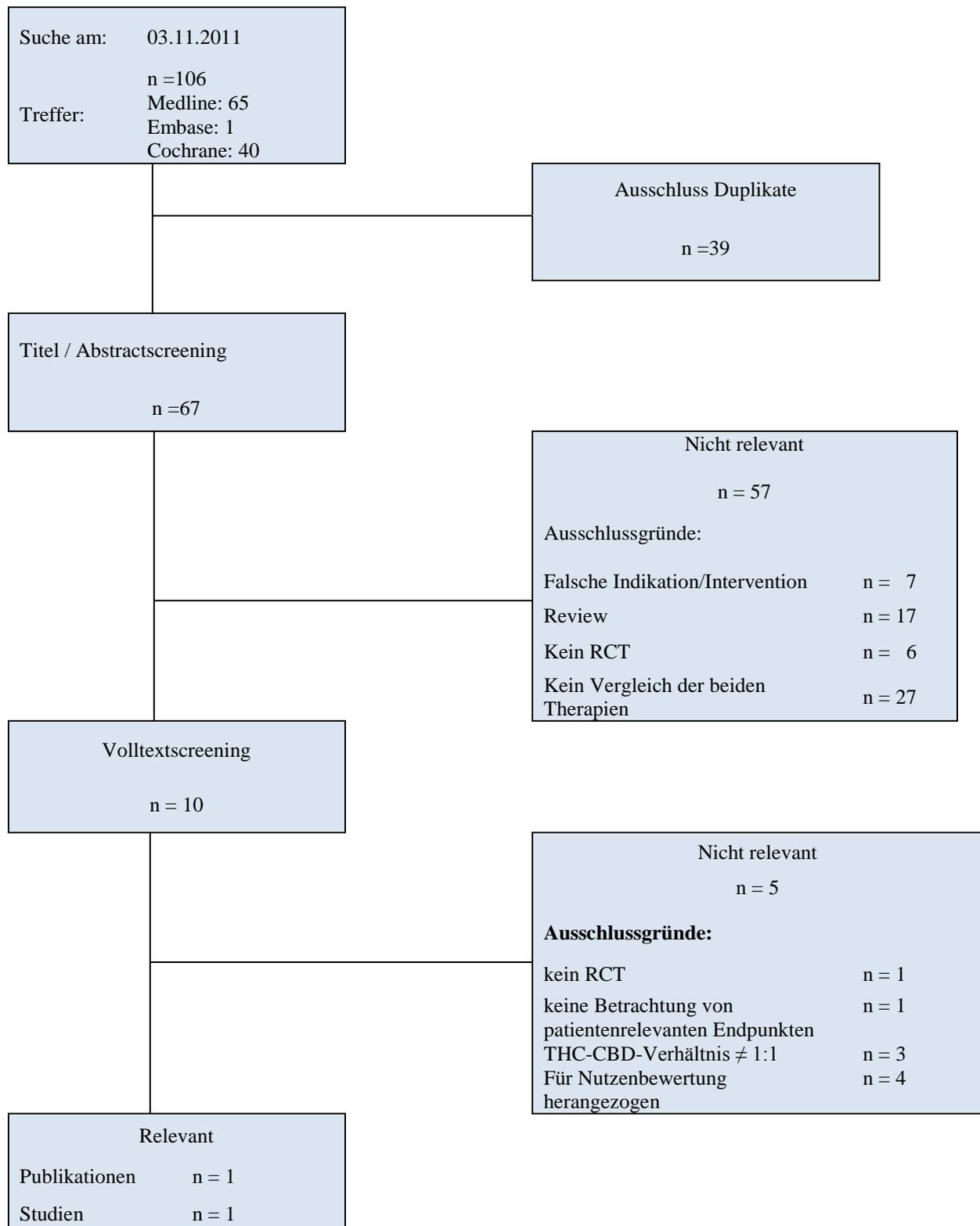


Abbildung 9: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach weiteren Untersuchungen mit dem zu bewertenden Arzneimittel

4.3.2.3.3 Studien aus der Suche in Studienregistern

Tabelle 4-65: Relevante Studien aus der Suche in Studienregistern – weitere Untersuchungen 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)
GWSP0702	clinicaltrials.gov [NCT00702468], (42) WHO-ICTRP [NCT00702468]	Ja	Ja (43)

a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.
b: n. d. = bibliografische Literaturrecherche nicht durchgeführt.

4.3.2.3.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

Tabelle 4-66: Studienpool – Weitere Untersuchungen 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])
placebokontrolliert						
GWSP0702	Nein	Ja	Nein	Ja (44)	Ja (42)	Ja (43)

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.

In den weiteren Untersuchungen wird die Studie GWSP0702 des Sponsors, die die Langzeitwirksamkeit überprüft, angeführt. Sie wurde in Tabelle 4-3 von der Nutzenbewertung für den direkten Vergleich ausgeschlossen, da unklar ist, ob die Patienten mittelschwere bis schwere Spastik aufwiesen. Allerdings erhielten alle an der Studie teilnehmenden Patienten im Vorfeld über einen längeren Zeitraum Sativex[®], so dass davon auszugehen ist, dass sie unter erheblicher Spastik gelitten haben. Diese Studie wurde von den Zulassungsbehörden als Ersatz einer Langzeitstudie zur Wirksamkeit und Sicherheit verlangt. Diese RCT hat als primären Endpunkt die Untersuchung des Wiederauftretens von Spastik nach Absetzen von

THC:CBD 1:1 bei Patienten mit MS, die dauerhaft auf eine Behandlung mit THC:CBD 1:1 angesprochen haben. Dabei wird der Grad der Spastizität auf der NRS nach dem Absetzen von Sativex® durch den Arzt bestimmt.

4.3.2.3.5 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.

Es handelt sich bei der Studie GWSP0702 um eine doppelblinde, 1:1 randomisierte, placebo-kontrollierte, parallelgruppierte Absetzstudie.

4.3.2.3.6 Ergebnisse aus weiteren Untersuchungen

4.3.2.3.6.1 <Endpunkt xxx> – weitere Untersuchungen

4.3.2.3.6.1.1 Absolute Veränderung des Grads der Spastizität auf der NRS– 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 4-67: Operationalisierung von Absolute Veränderung des Grads der Spastizität auf der NRS

Studie	Operationalisierung
GWSP0702	Es handelt sich bei dieser Studie um eine Studie zur Überprüfung der Langzeitwirksamkeit, sodass für die randomisierte Behandlungsphase, zumindest unter Placebo, eine Zunahme des Schweregrads der Spastik zu erwarten war, weil eine bestehende Therapie mit THC:CBD 1:1 durch eine Placebo-Therapie ersetzt wurde.. Der Schweregrad der Spastik wurde von den Patienten im Tagebuch dokumentiert über die gesamte Studiendauer (4 Wochen nach Absetzen oder bis zum vorzeitigen Studienende). Im Tagebuch wurde die durchschnittliche Schwere der Spastik über die vergangenen 24 Stunden auf einer 11-stufigen NRS (0-10) dokumentiert (0= 'no spasticity', 10= 'worst ever spasticity'). Der absolute Rückgang der Schwere der Spastik wurde patientenbezogen analysiert als Änderung des 7 Tagedurchschnitts des NRS zum Studienende im Vergleich zu Baseline.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind,

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

Tabelle 4-68: Bewertung des Verzerrungspotenzials für Absolute Veränderung des Grads der Spastizität auf der NRS 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
GWSP0702	Niedrig	Ja	Ja	Nein	Nein	Niedrig

In der Studie war der Endpunkterheber (Patient) verblindet. Der Endpunkt wurde wie geplant für die ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf eine ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt, wie geplant, statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in der Studie das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

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.

Tabelle 4-69: Ergebnisse für Absolute Veränderung des Grads der Spastizität auf der NRS aus RCT mit dem zu bewertenden Arzneimittel

Studie	Veränderung des Grads der Spastizität (NRS Score)								
	Baseline			Studienende			Veränderung		
	N	Mean (SD)	Median Min-Max	N	Mean (SD)	Median Min-Max	Mean (SD)	Median Min-Max	Gruppen-differenz (*) (95%-CI) p-Wert
GWSP 0702									
Sativex®	18	3,60 (1,67)	3,29 (1,3-6,3)	18	4,71 (1,99)	4,50 (1,1-8,3)	1,11	0,79	-0,21 (-1,22; 0,79)
Placebo	18	4,13 (2,23)	4,43 (0,0-8,0)	18	5,24 (2,17)	5,62 (1,0-8,7)	1,10	1,20	p=0,7204
(*) Adjustiertes lineares Modell Veränderung 4 Wochen nach Fortführung versus Absetzen nach dauerhaftem Ansprechen auf Behandlung mit Sativex®									

Im direkten Vergleich zur Therapiefortsetzung ohne die Wirkstoffkombination THC:CBD 1:1 zeigt die Add-on Therapie unter THC:CBD 1:1 eine positive Gruppendifferenz auf der 11-stufigen NRS, die aber statistisch nicht signifikant ist.

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

4.3.2.3.6.1.2 Schlafqualität– 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 4-70: Operationalisierung von Schlafqualität

Studie	Operationalisierung
GWSP0702	Die Schlafqualität wurde als Unterbrechung des Schlafs von den Patienten im Tagebuch täglich dokumentiert über die gesamte Studiendauer bis 4 Wochen nach randomisierter Fortsetzung versus Absetzen der Studienmedikation oder bis zum vorzeitigen Studienende. Die Dokumentation der Schlafqualität erfolgte auf einer 11-stufigen NRS (0= 'No Sleep Disruption' - 10= 'Completely disrupted (unable to sleep at all)'). Die Auswertung erfolgte als Differenz der Beurteilung im Vergleich zu Baseline (Randomisierung in Woche 0).

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

Tabelle 4-71: Bewertung des Verzerrungspotenzials für Schlafqualität 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
GWSP0702	Niedrig	Ja	Ja	Nein	Nein	Niedrig

In der Studie war der Endpunkterheber (Patient) verblindet. Der Endpunkt wurde, wie geplant, für die ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf eine ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt, wie geplant, statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in der Studie das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

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.

Tabelle 4-72: Ergebnisse für Schlafqualität aus RCT mit dem zu bewertenden Arzneimittel

Studie	Schlafqualität (NRS 0-10 Beurteilung) Erhoben als Schlafunterbrechung GWMS0702)								
	Baseline			Studienende			Veränderung		
	N	Mean (SD)	Median Min-Max	N	Mean (SD)	Median Min-Max	Mean (SD)	Median Min-Max	Gruppen- lifferenz (*) (95%-CI) p-Wert
GWMS 0702									
Sativex®	18	2,81 (1,71)	1,93 (0,0-5,7)	18	2,9 (2,48)	2,36 (0,0-7,7)	0,60 (1,84)	0,36 (-2,4-5,7)	-0,64 (-1,6; 0,33) p=0,2710
Placebo	18	2,11 (1,75)	2,00 (0,0-6,9)	18	3,36 (2,18)	3,0 (0,1-8,0)	1,25 (1,52)	0,93 (-1,7-4,9)	
(*)lineares Modell Verbesserung 4 Wochen nach Fortführung versus Absetzen nach dauerhaftem Ansprechen (mindestens 20% Verbesserung der NRS) auf Sativex®									

Im direkten Vergleich zur Therapiefortsetzung ohne Sativex® zeigt die Add-on Therapie unter Sativex® eine Besserung der Schlafqualität. Die Daten der vor Erreichen der vorgesehenen Fallzahl beendeten Absetzstudie unter Langzeitnutzern von Sativex® zeigte eine Differenz von -0,64, die statistisch nicht signifikant ist (p=0,2710), wahrscheinlich bedingt durch die geringe Zahl von Patienten.

4.3.2.3.6.1.3 Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC)– 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 4-73: Operationalisierung von Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC)

Studie	Operationalisierung
GWSP0702	Die Gesamtbeurteilung wurde auf Global Impression of Change Skalen am Studienende (Woche 4) vom Patienten (SGIC) und Caregiver (CGIC) dokumentiert. Die GIC wurde erfasst als Beurteilung der Verbesserung des Gesamtzustands auf 7-Punkt Skalen (1 = <i>Very Much Improved</i> , 2 = <i>Much Improved</i> , 3 = <i>Minimally Improved</i> , 4 = <i>No Change</i> , 5 = <i>Minimally Worse</i> , 6 = <i>Much Worse</i> , 7 = <i>Very Much Worse</i>). Zusätzlich zur allgemeinen funktionellen Globalbeurteilung der klinischen Veränderung wurde vom Betreuer die Änderung der Transferierbarkeit auf der GIC bestimmt.

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

Tabelle 4-74: Bewertung des Verzerrungspotenzials für Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC) 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
GWSP0702	Niedrig	Ja	Ja	Nein	Nein	Niedrig

In der Studie war der Endpunkterheber (Patient) verblindet. Der Endpunkt wurde wie geplant für die ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf eine ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt wie geplant statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in der Studie das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

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.

Tabelle 4-75: Ergebnisse für Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC) aus RCT mit dem zu bewertenden Arzneimittel

Studie	Beurteilung der klinischen Änderung am Studienende durch den Patienten								Odds ratio (95%-CI) p-Wert
	N	Sehr viel besser	Viel besser	Wenig besser	Keine Änderung	Wenig schlechter	Viel schlechter	Sehr viel schlechter	
GWMS 0702									
Sativex®	18	0	0	0	6 (33)	6 (33)	5 (28)	1 (6)	4,55 (1,58, 13,99)
Placebo	18	0	0	0	1 (6)	5 (28)	9 (50)	3 (17)	p=0,0172
Verbesserung 4 Wochen nach Fortführung versus Absetzen nach dauerhaftem Ansprechen (mindestens 20 % Verbesserung der NRS) auf Sativex®									

Im direkten Vergleich zur Therapiefortsetzung ohne die Wirkstoffkombination THC:CBD 1:1 zeigt die Therapiefortsetzung mit THC:CBD 1:1 eine bessere abschließende Beurteilung der globalen Veränderung durch die Patienten. Die resultierende Odds ratio betrug 4,55 zugunsten von THC:CBD 1:1. Das Ergebnis ist mit $p=0,0172$ statistisch signifikant.

Die Odds ratio wurde mittels der logistischen Regressionstechnik berechnet.

4.3.2.3.6.1.4 Globalbeurteilung der klinischen Veränderung durch die betreuende Person (CGIC)– 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 4-76: Operationalisierung von klinische Veränderung durch die betreuende Person (CGIC)

Studie	Operationalisierung
GWSP0702	Die Gesamtbeurteilung wurde auf Global Impression of Change Skalen am Studienende (Woche 4) vom Patienten (SGIC) und Caregiver (CGIC) dokumentiert. Die GIC wurde erfasst als Beurteilung der Verbesserung des Gesamtzustands auf 7-Punkt Skalen (1 = <i>Very Much Improved</i> , 2 = <i>Much Improved</i> , 3 = <i>Minimally Improved</i> , 4 = <i>No Change</i> , 5 = <i>Minimally Worse</i> , 6 = <i>Much Worse</i> , 7 = <i>Very Much Worse</i>).

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

Tabelle 4-77: Bewertung des Verzerrungspotenzials für Globalbeurteilung der klinischen Veränderung durch die betreuende Person (CGIC) 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
GWSP0702	Niedrig	Ja	Ja	Nein	Nein	Niedrig

In der Studie war der Endpunkterheber verblindet. Der Endpunkt wurde wie geplant für die ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf eine ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt wie geplant statistisch ausgewertet wurde und die Ergebnisse im

vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in der Studie das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

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.

Tabelle 4-78: Ergebnisse für Globalbeurteilung der klinischen Veränderung durch die betreuende Person (CGIC) aus RCT mit dem zu bewertenden Arzneimittel

Studie	Beurteilung der klinischen Änderung am Studienende durch die betreuende Person								
	Funktionale Möglichkeiten des Patienten								
Anzahl Patienten n (%)									
	N	Sehr viel besser	Viel besser	Wenig besser	Keine Änderung	Wenig schlechter	Viel Schlechter	Sehr viel Schlechter	Odds ratio (95%-CI) p-Wert
GWSP 0702									
Sativex®	10	0	0	0	3 (30)	5 (50)	2 (20)	0	18,55 (3,94; 118,77)
Placebo	14	0	0	0	0	3 (21)	9 (64)	2 (14)	p=0,0011
Transferierbarkeit des Patienten									
Anzahl Patienten n (%)									
GWSP 0702									
Sativex®	10	0	0	0	3 (30)	5 (50)	2 (20)	0	3,444 (0,948; 13,718)
Placebo	14	0	0	0	2 (14)	5 (36)	6 (43)	1 (7)	p=0,1151
Verbesserung 4 Wochen nach Fortführung versus Absetzen nach dauerhaftem Ansprechen (mindestens 20% Verbesserung der NRS) auf Sativex®									

Im direkten Vergleich zur Therapiefortsetzung ohne die Wirkstoffkombination THC:CBD 1:1 zeigt die Therapiefortsetzung mit THC:CBD 1:1 eine bessere abschließende Beurteilung der globalen Veränderung der funktionalen Möglichkeiten des Patienten durch die betreuende Person. Die resultierende Odds ratio betrug 18,55 zugunsten von THC:CBD 1:1. Das Ergebnis ist mit $p=0,0011$ statistisch signifikant. Die abschließende Beurteilung der globalen Veränderung der Transferierbarkeit des Patienten ergab ebenfalls eine bessere abschließende Beurteilung durch die betreuenden Personen bei Therapiefortsetzung unter THC:CBD 1:1 im direkten Vergleich zur Therapiefortsetzung ohne die Add-on Therapie. Die resultierende Odds ratio betrug in diesem Fall 3,444 zugunsten von THC:CBD 1:1, ist aber mit $p=0,1151$ statistisch nicht signifikant. Zu beachten ist, dass in der Studie zur Untersuchung der Langzeitwirksamkeit GWSP0702 der Effekt nach Fortsetzen (Sativex®-Gruppe) bzw. nach

Absetzen (Placebogruppe) der Therapie für Patienten beurteilt wurde, die vorher positiv auf die Behandlung mit Sativex® angesprochen haben.

4.3.2.3.6.1.5 Gehfähigkeit – 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 4-79: Operationalisierung von Gehfähigkeit

Studie	Operationalisierung
GWSP0702	Die 10 m-Gehstrecke wurde als Zeit für das Zurücklegen einer geraden Strecke von 10m durch den Investigator erfasst.

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

Tabelle 4-80: Bewertung des Verzerrungspotenzials für Gehfähigkeit (10 m-Gehstreckentest) 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
GWSP0702	Niedrig	Ja	Ja	Nein	Nein	Niedrig

In der Studie war der Endpunkterheber verblindet. Der Endpunkt wurde wie geplant für die ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf eine ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt, wie geplant, statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in der Studie das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

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.

Tabelle 4-81: Ergebnisse für Gehfähigkeit (10 m-Gehstreckentest) aus RCT mit dem zu bewertenden Arzneimittel

Studie	10 m-Gehstreckentest (Sekunden)								
	Baseline			Studienende			Veränderung		
	N	Mean (SD)	Median Min-Max	N	Mean (SD)	Median Min-Max	Mean (SD)	Median Min-Max	Gruppen- differenz (*) (95%-CI) p-Wert
GWSP 0702									
Sativex®	12	40,1 (41,78)	19,5 (10-141)	11	47,3 (50,29)	21 (12-171)	4,5 (13,92)	2 (-20-20)	-1,78 (-14,52; 10,96)
Placebo	9	24,1 (17,19)	17 (8-56)	4	18,3 (4,5)	19 (13-22)	2,5 (2,65)	3 (-1-5)	p=0,8076(**)
(*)GWMS0702: lineares Modell (**) Die adjustierten Mittelwerte (LSMeans) betragen 3,46 in der Sativex® Gruppe und 5,24 in der Placebo Gruppe Veränderung 4 Wochen nach Fortführung versus Absetzen nach dauerhaftem Ansprechen (mindestens 20 % Verbesserung der NRS) auf Sativex®									

In der Studie ergab sich eine nichtsignifikante Differenz von -1,78. Die Geschwindigkeitsänderung für den Behandlungsarm THC:CBD 1:1 liegt mit -0,038 m/s (10 m/47,3 s am Studienende minus 10 m/40,1 s) über der klinisch relevanten Differenz von 0,05 m/s liegt.

4.3.2.3.6.1.6 Nebenwirkungen

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

Tabelle 4-82: Operationalisierung von Nebenwirkungen

Studie	Operationalisierung
GWSP0702	Nebenwirkungen wurden über den gesamten Studienzeitraum GCP-konform dokumentiert bis einschließlich der Follow-up Periode nach Studienende nach Anzahl der Patienten.

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

Tabelle 4-83: Bewertung des Verzerrungspotenzials für Nebenwirkungen 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
GWSP0702	Niedrig	Ja	Ja	Nein	Nein	Niedrig

In der Studie war der Endpunkterheber verblindet. Der Endpunkt wurde, wie geplant, für die ITT-Population ausgewertet, die in einem Flow-Chart hinsichtlich von Lost-to-Follow-ups dargestellt sind. Hinweise auf eine ergebnisgesteuerte Darstellung des Endpunkts liegen nicht vor, da der Endpunkt, wie geplant, statistisch ausgewertet wurde und die Ergebnisse im vorliegenden Clinical Study Report berichtet sind. Entsprechend kann in der Studie das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft werden.

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.

Tabelle 4-84: Ergebnisse für Nebenwirkungen aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anzahl Patienten mit Nebenwirkungen				Anzahl der Nebenwirkungen	
	Sativex [®]		Placebo		Sativex [®]	Placebo
	N	n (%)	N	n (%)		
GWSP0702	18	15 (83)	18	14 (78)	25	26

Die Ergebnisse zeigen, dass die Zahl der Patienten mit Nebenwirkungen in beiden Behandlungsarmen ähnlich ist. Dies lässt aber keine eindeutige Aussage zur Sicherheit des Arzneimittels zu. Deshalb wurde eine weitere Analyse zur Häufigkeit des Auftretens von schwerwiegenden Nebenwirkungen mit möglichem Zusammenhang zur Studienmedikation durchgeführt. Diese Ergebnisse zeigen, dass in dieser Studie keine schwerwiegenden Nebenwirkungen im Zusammenhang mit Sativex[®] aufgetreten sind (Tabelle 4-85).

Tabelle 4-85: Ergebnisse für schwerwiegende Nebenwirkungen aus RCT mit dem zu bewertenden Arzneimittel

Studie	Anzahl Patienten mit schwerwiegenden Nebenwirkungen			
	Sativex [®]		Placebo	
	N	n (%)	N	n (%)
GWSP0702	18	0 (0)	18	0 (0)

4.3.2.3.6.2 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.

Es wurden keine Subgruppenanalysen durchgeführt.

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.

In folgender Tabelle sind die Ergebnisse der einzelnen Endpunkte zusammengefasst dargestellt (Tabelle 4-86):

Tabelle 4-86: Zusammenfassende Einschätzung aller Endpunkte für die Studie GWSP0702

Endpunkt	
Verbesserung des Schweregrads der Spastik	+
Schlafqualität	+
Gehfähigkeit	±
Klinische Globalbeurteilung	±
SGIC	++
CGIC (allgemeine)	++
CGIC (Transferierbarkeit)	+
Nebenwirkungen	±
++ statistisch signifikanter Vorteil von THC:CBD 1:1 gegenüber Placebo + numerischer Vorteil von THC:CBD 1:1 gegenüber Placebo ± kein Vorteil von THC:CBD 1:1 gegenüber Placebo	

Betrachtet man die funktionellen patientenrelevanten Endpunkte klinische Globalbeurteilung und die Verbesserung des Schweregrads der Spastik weist die Add-on Therapie mit der Wirkstoffkombination THC:CBD 1:1 gegenüber der Vergleichstherapie einen klaren Vorteil auf.

In der Placebogruppe kam es zu überdurchschnittlich vielen Studienabbrüchen bedingt durch Unzufriedenheit mit der Medikation und anschließender Rückkehr zur vorherigen Behandlung mit Sativex®.

Da keine Meta-Analysen durchgeführt wurden, liegen keine numerischen Ergebnisse vor.

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

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

Die im Dossier dargestellten Ergebnisse der einzelnen RCT stellen Evidenz der Stufe I dar. Alle vorgelegten Studien waren randomisiert, kontrolliert, doppelblind angelegt und zu allen Erhebungszeitpunkten bestand eine Verblindung der Endpunkterheber für die verschiedenen Endpunkte. Wie dargestellt, bestand basierend auf den dargestellten Kriterien ein niedriges Verzerrungspotential sowohl auf Studienebene, als auch auf Endpunktebene. Alle vorgelegten RCT können als methodisch hochwertige Einzelstudien eingestuft werden.

Die zur Darstellung des Zusatznutzens im Dossier vorgelegte Evidenz basiert auf den im Rahmen der Zulassung vorgelegten klinischen Abschlussberichten zu vier der Studien (GWMS0001, GWMS0106, GWCL0403 und GWSP0604) sowie der Publikation von Ergebnissen zu einer fünften Studie (ARA 2009). Alle fünf RCT berichten Ergebnisse zu patientenrelevanten, methodisch validierten Endpunkten, die für die Zusatznutzenbewertung von Sativex[®] relevant sind, und wurden im Dossier entsprechend berücksichtigt. Die Studie GWSP0702 wurde nur in den weiteren Unterlagen aufgeführt, da sie nicht zu 100 % den Einschlusskriterien entsprach, aber als von den Zulassungsbehörden geforderte Studie, die das Absetzen von Sativex[®] untersucht und somit die Langzeitwirksamkeit von Sativex[®] überprüft, unseres Erachtens zu berücksichtigen ist.

Zwei der fünf Studien (GWCL0403 und GWSP0604) waren konfirmatorische Studien zum Testen statistischer Hypothesen und in der Fallzahl hinsichtlich des primären Endpunkts gepowert. Drei weitere Studien (GWMS0001, GWMS0106 und ARA2009) hatten eine explorative Zielsetzung und eine statistische Testung von Hypothesen wurde nicht durchgeführt.

Zwei der Studien (GWSP0604 und ARA2009) untersuchten die dauerhafte Zusatzbehandlung mit THC:CBD 1:1 unter Fortführung der anti-spastische Arzneimitteltherapie bei Patienten, die auf einen Therapieversuch mit Sativex[®] bereits positiv angesprochen hatten. Entsprechend zeigen die Ergebnisse aus diesen Studien den Zusatznutzen der Zusatztherapie mit THC:CBD 1:1 explizit für die Patientenpopulation, die der zugelassenen Indikationsstellung für Sativex[®] entspricht (Patienten, die auf einen 4-wöchigen Therapieversuch mit Sativex[®] positiv angesprochen haben). Unter diesen zwei Studien liefert die Studie GWSP0604 aufgrund ihrer konfirmatorischen Anlage die stärkste Evidenz im Vergleich zu der explorativ angelegten Studie ARA2009 mit vergleichsweise geringeren Fallzahlen. Außerdem wurden in der Studie GWSP0604 nur die Ergebnisse von Patienten ausgewertet, die auf eine Therapie mit Sativex[®] ansprechen. Somit bildet diese Population die Indikation der Zulassung und den tatsächlichen Einsatz am besten ab.

Alle dem Dossier zugrundeliegenden RCT untersuchten patientenrelevante Endpunkte unter der Zusatzbehandlung mit der Wirkstoffkombination THC:CBD 1:1 unter Fortführung der anti-spastischen Arzneimitteltherapie, woraus sich direkte Aussagen zum Zusatznutzen von THC:CBD 1:1 zur Behandlung der Spastik bei Patienten mit MS ableiten lassen. Die zwei Studien GWSP0604 und ARA2009 untersuchten dabei die Zusatzbehandlung mit THC:CBD 1:1 unter Fortführung der anti-spastischen Arzneimitteltherapie für die Patientenpopulation, die der jetzt zugelassenen Indikationsstellung von Sativex[®] explizit entspricht.

Beleg für den Zusatznutzen von Sativex[®] liefern Ergebnisse aus drei RCT, die einen statistisch signifikanten Unterschied zugunsten von Sativex[®] in der Verbesserung des Grads der Spastik auf der NRS Skala zeigen (GWMS0001, GWMS0106, GWSP0604). Ergänzende Hinweise ergeben sich aus einem weiteren RCT, der für diesen Endpunkt ebenfalls einen Unterschied zugunsten von THC:CBD 1:1 zeigte, die statistisch nicht signifikant waren (GWCL0403) und den Ergebnissen einer Meta-Analyse aus den Ergebnissen zu den drei RCT GWMS0001, GWMS0106 und GWCL0403 (Abbildung 4).

Für die relevant verbesserte Rate von Patienten mit einer klinisch relevanten Verbesserung der Spastik (Verbesserung um mindestens 30 % auf der NRS) ergeben sich gleichfalls Belege aus zwei der RCT, die einen Gruppenunterschied zugunsten von THC:CBD 1:1 zeigten (GWMS0106, GWSP0604). Weitere Hinweise liegen aus den Ergebnissen einer dritten RCT vor, die diesen Endpunkt ebenfalls untersuchte, und Gruppenunterschiede zugunsten von THC:CBD 1:1 zeigte, die nicht statistisch signifikant waren (GWCL0403) sowie aus den Ergebnissen einer Meta-Analyse aus den Ergebnissen zu den drei RCT GWMS0001, GWSP0106 und GWCL0403 (Abbildung 5).

Weitere Hinweise zum Zusatznutzen der Sativex[®] Therapie ergeben sich aus Ergebnissen der RCT zu folgenden anderen patientenrelevanten Endpunkten:

- Die Anzahl der Spastiken pro Tag zeigte sich in den zwei RCT, die diesen Endpunkt untersuchten, besser unter THC:CBD 1:1. In einer RCT war der Gruppenunterschied statistisch signifikant (GWSP0604), in der anderen nicht (GWMS0106).
- Die Schlafqualität unter THC:CBD 1:1 war in den drei RCT, die diesen Endpunkt untersuchten, besser unter THC:CBD 1:1. In einer RCT war der Gruppenunterschied statistisch signifikant (GWSP0604), in den beiden anderen nicht (GWCL0403, GWMS0106).
- Die Globalbeurteilung der Veränderung durch den Patienten (SGIC Beurteilung) war in beiden RCT, die diesen Endpunkt untersuchten, besser unter THC:CBD 1:1. In einer RCT war der Gruppenunterschied statistisch signifikant (GWSP0604), in der anderen RCT nicht (GWMS0106).
- Die Globalbeurteilung der Veränderung der funktionalen Möglichkeiten des Patienten durch die betreuende Person (CGIC Beurteilung) war in den beiden RCT, die diesen Endpunkt untersuchten, besser unter THC:CBD 1:1. In einem RCT war der Gruppenunterschied statistisch signifikant (GWSP0604), in einem zweiten RCT nicht (GWCL0403). Die Beurteilung der Transferierbarkeit des Patienten war in einem RCT besser unter THC:CBD 1:1, statistisch aber nicht signifikant.
- Die Globalbeurteilung der Veränderung der funktionalen Möglichkeiten des Patienten durch den Arzt (PhGIC Beurteilung) war in einer RCT, die diesen Endpunkt untersuchte, besser unter THC:CBD 1:1. Der Gruppenunterschied war statistisch signifikant (GWSP0604).

- Die Aktivitäten des täglichen Lebens (Barthel Index) waren in den beiden RCT, die diesen Endpunkt untersuchten, besser unter THC:CBD 1:1. In einer der RCT war der Gruppenunterschied statistisch signifikant (GWSP0604), in der zweiten RCT nicht (GWCL0403).
- Die 10 m-Gehstrecke war in den zwei RCT, die diesen Endpunkt untersuchten, besser unter THC:CBD 1:1. In den beiden RCT war der Gruppenunterschied nicht statistisch signifikant auf dem 5 % Niveau (GWCL0403, GWSP0604).
- Der Gesundheitszustand unter THC:CBD 1:1 war in zwei RCT, die diesen Endpunkt untersuchten, leicht verbessert unter THC:CBD 1:1. In beiden Studien war der Gruppenunterschied nicht statistisch signifikant (GWCL0403, GWSP0604).

Aus allen fünf RCT zeigte sich ein gutes Sicherheitsprofil von Sativex® ohne relevante Nebenwirkungen durch die Kombination zweier Cannabinoide.

Der Zusatznutzen wurde mit Hilfe der Auswertung zahlreicher patientenrelevanter Endpunkte bestimmt. Dazu gehören:

1. Verbesserung des Grads der Spastik
2. Anzahl der Patienten mit einem klinisch relevanter Rückgang der Schwere der Spastik
3. Häufigkeit der Spasmen pro Tag
4. Globale Beurteilung der Veränderung der Spastik durch den Patienten, die betreuende Person und den Arzt
5. Gesundheitszustand
6. Lebensqualität
7. Schlafqualität
8. Aktivitäten des täglichen Lebens
9. Gehfähigkeit
10. Verringerung der Nebenwirkungen

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):

- 1. erheblicher Zusatznutzen*
- 2. beträchtlicher Zusatznutzen*
- 3. geringer Zusatznutzen*
- 4. nicht quantifizierbarer Zusatznutzen*
- 5. kein Zusatznutzen belegbar*
- 6. 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.

Aus den gezeigten und analysierten patientenrelevanten Endpunkten lässt sich eine wesentliche Abschwächung schwerwiegender funktioneller Symptome wie z.B. der Spastik (primärer Endpunkt) sowie die Verbesserung in den Aktivitäten des täglichen Lebens (Barthel-Index, sekundärer Endpunkt) beobachten. Dieses Ergebnis wird unterstützt durch die verbesserte Globalbeurteilung der Veränderung der funktionalen Möglichkeiten durch den Arzt. Nach §5 (7) Nr. 2 des 5. Kapitels der VerFO wurde gegenüber der Vergleichstherapie eine bisher nicht erreichte deutliche Verbesserung des therapielevanten Nutzens durch eine Abschwächung schwerwiegender Symptome gezeigt, so dass ein beträchtlicher Zusatznutzen vorliegt. Dieser wurde anhand von zahlreichen patientenrelevanten Endpunkten belegt.

Der Zusatznutzen der Zusatztherapie mit Sativex[®] Spray zur Anwendung in der Mundhöhle von Patienten mit mittelschwerer bis schwerer Spastik aufgrund von Multipler Sklerose (MS), die nicht angemessen auf eine andere anti-spastische Arzneimitteltherapie angesprochen haben und die eine klinisch erhebliche Verbesserung von mit der Spastik verbundenen Symptomen während eines Anfangstherapieversuchs aufzeigen, ist beträchtlich. Er zeigt sich in einer statistisch signifikant höheren Anzahl an Patienten, die eine klinisch relevante Verbesserung ihres Grades der Spastik anhand der NRS erfahren sowie in einem statistisch signifikanten und klinisch relevanten Rückgang der durchschnittlichen Anzahl von Spastiken pro Tag bei diesen Patienten. Patienten in der für die Behandlung mit Sativex[®] zugelassenen Population zeigen unter der Wirkstoffkombination THC:CBD 1:1 zudem im Gruppenvergleich eine statistisch signifikant ausgeprägtere Verbesserung der Schlafqualität sowie eine statistisch signifikant stärkere Verbesserung des Barthel Index als Maß für die Aktivitäten des

täglichen Lebens sowie eine größere Verbesserung in ihrer Gehfähigkeit und Lebensqualität. Entsprechend wurde die Zusatzbehandlung mit THC:CBD 1:1 im Vergleich statistisch signifikant global um mindestens eine Kategorie besser beurteilt sowohl durch die Patienten selbst, als auch durch die betreuenden Personen und behandelnden Ärzte.

Die Ergebnisse aller der dem Dossier zur Zusatznutzenbewertung zugrunde liegenden RCT zeigen einen Gruppenvorteil der Zusatztherapie mit Sativex[®], basierend auf den dargestellten patientenrelevanten Endpunkten. Eine im Gruppenvergleich klinisch relevante Verbesserung in der THC:CBD 1:1-Behandlungsgruppe war für alle betrachteten Studienpopulationen feststellbar, d.h. auch für gemischte Populationen von Patienten, die auf eine Therapie mit Sativex[®] prinzipiell ansprechen und nicht ansprechen (Studien GWMS0001, GWMS0106 und GWCL0403). In den Studien mit Patientenpopulationen, die der zugelassenen Indikation von Sativex[®] entsprachen (Studien GWSP0604 und ARA2009), d.h. in Studien mit Patienten, die in einem Therapieversuch auf Sativex[®] angesprochen haben, war der Effekt der Zusatztherapie mit THC:CBD 1:1 im Gruppenvergleich nochmals deutlicher ausgeprägt. In der für diese Population konfirmatorischen Studie (GWSP0604) war die Gruppendifferenz für den primären Endpunkt, sowie für nahezu alle betrachteten sekundären Endpunkte statistisch signifikant auf dem 5 % Niveau. Unterstützend zeigten die explorative Studie ARA2009, die an Patientenpopulationen durchgeführt wurden, die der zugelassenen Indikation von Sativex[®] entsprachen, bereits anhand von sehr geringen Fallzahlen ähnlich deutliche Tendenz im Gruppenvergleich der betrachteten Endpunkte zugunsten der Zusatztherapie mit THC:CBD 1:1.

Der beträchtliche Zusatznutzen einer Zusatztherapie mit Sativex[®] von Patienten mit mittelschwerer bis schwerer Spastik aufgrund von Multipler Sklerose (MS), die nicht angemessen auf eine andere anti-spastische Arzneimitteltherapie angesprochen haben und die eine klinisch erhebliche Verbesserung von mit der Spastik verbundenen Symptomen während eines Anfangstherapieversuchs aufzeigen, ist in dem hier vorliegenden Dossier belegt. Die im Dossier als Resultat der systematischen Übersicht dargelegten Nachweise zum Zusatznutzen einer Zusatztherapie mit Sativex[®] berücksichtigen Ergebnisse von fünf verschiedenen RCT und entsprechen so den Anforderungen des Evidenzlevels der Stufe 1.

Für die systematische Identifizierung der dem Dossier zugrunde liegenden Ergebnisse wurden öffentlich zugängliche Quellen zu klinischen Studien, Berichte zu klinischen Studien und Literaturdatenbanken wie im Dossier dargestellt, sowie die beim pharmazeutischen Hersteller verfügbaren Informationen zu relevanten klinischen Studien voll umfänglich berücksichtigt. Insofern liegt Ergebnisvollständigkeit vor. Alle fünf als relevant für die Zusatznutzenbewertung von Sativex[®] identifizierten Studien sind RCT und entsprechen damit der Evidenzklasse 1. Hinweise auf ein relevantes Verzerrungspotential dieser RCT liegen nicht vor. Drei der fünf RCT wurden an Patientenpopulationen durchgeführt, die der jetzt zugelassenen Anwendung von Sativex[®] in Deutschland entsprechen hinsichtlich einer unzureichenden Wirkung der Vortherapie der Spastik. Eine der RCT wurde konfirmatorisch mit der für die statistische Power erforderlichen Fallzahl durchgeführt. Insofern kann, basierend auf Ergebnisvollständigkeit der dargestellten Endpunkte, von der hinreichenden

Sicherheit des dargestellten Zusatznutzens von Sativex® in der in Deutschland zugelassenen Anwendung ausgegangen werden.

Das Ausmaß des Zusatznutzens der Zusatztherapie mit Sativex® bei gemäß Zulassung für die Behandlung in Frage kommenden Patienten ist beträchtlich. Unter Zusatztherapie mit Sativex® wird für Patienten mit mittelschwerer bis schwerer Spastik aufgrund von Multipler Sklerose (MS), die nicht angemessen auf eine andere anti-spastische Arzneimitteltherapie angesprochen haben und die eine klinisch erhebliche Verbesserung von mit der Spastik verbundenen Symptomen während eines Anfangstherapieversuchs aufzeigen, ein bisher nicht erreichter patientenrelevanter therapeutischer Zusatzeffekt erzielt. Unter Sativex® wird eine gegenüber der zweckmäßigen Vergleichstherapie deutliche Verbesserung des Gesundheitszustands der Patienten erreicht, vor allem durch die unter THC:CBD 1:1 häufiger eintretende klinisch relevante Abnahme des Grads der Spastik sowie die deutliche Abnahme der durchschnittlichen Anzahl der Spastiken pro Tag. Mehr Patienten nehmen unter Zusatztherapie mit Sativex® insgesamt eine spürbare Linderung der Erkrankung wahr, was sowohl von den Patienten selbst als auch durch die den Patienten betreuenden oder behandelnden Personen berichtet wird.

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 4-87: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Patienten mit mittelschwerer bis schwerer Spastik aufgrund von Multipler Sklerose (MS), die während eines 4-wöchigen Therapieversuchs auf die Zusatzbehandlung mit Sativex® ansprechen	beträchtlich

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.

ND – Bei Sativex[®] handelt es sich nicht um ein Arzneimittel zur Behandlung von seltenen Leiden.

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 4-88: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

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

ND – Es wurde kein indirekter Vergleich vorgenommen.

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.

ND – Es wurde kein indirekter Vergleich vorgenommen.

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.

ND – Es lagen valide Daten zu den patientenrelevanten Endpunkten vor.

4.5.4 Verwendung von Surrogatendpunkten

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

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005¹³, Molenberghs 2010¹⁴). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

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

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

¹⁴Molenberghs 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.

(Burzykowski 2006¹⁵) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006¹⁶) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

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

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

Sofern Sie im Dossier Ergebnisse zu Surrogatendpunkten eingereicht haben, benennen Sie die Gründe für die Verwendung von Surrogatendpunkten. Beschreiben Sie, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

ND – Im Dossier wurden keine Ergebnisse zu Surrogatendpunkten eingereicht.

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.

- GWMS0001: Clinical Study Report of study code GWMS0001
- GWMS0106: Clinical Study Report of study code GWMS0106

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

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

- GWCL0403: Clinical Study Report of study code GWCL0403
- GWSP0604: Clinical Study Report of study code GWSP0604
- GWSP0702: Clinical Study Report of study code GWSP0702
- ARA 2009: Aragona et al. 2009

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. Clark P, Altman D. Assessment of Blinding in Phase III Sativex® spasticity studies. . In: Pharmaceuticals G, editor. 2006.
2. Farrar JT, Troxel AB, Stott C, Duncombe P, Jensen MP. Validity, reliability, and clinical importance of change in a 0-10 numeric rating scale measure of spasticity: a post hoc analysis of a randomized, double-blind, placebo-controlled trial. *Clinical therapeutics*. 2008;30(5):974-85. Epub 2008/06/17.
3. Wade DT, Collin C, Stott C, Duncombe P. Meta-analysis of the efficacy and safety of Sativex (nabiximols), on spasticity in people with multiple sclerosis. *Mult Scler*. 2010;16(6):707-14. Epub 2010/06/19.
4. Pharmaceuticals G. Clinical Study Report GWMS0001: A double blind, randomised, parallel group, placebo controlled trial of a combination of delta-9-tetrahydrocannabinol (THC) and cannabidiol (CBD) in patients with multiple sclerosis, followed by an open label assessment and study extension. 2006 NCT00202423.
5. Lechner HE, Frotzler A, Eser P. Relationship between self- and clinically rated spasticity in spinal cord injury. *Archives of physical medicine and rehabilitation*. 2006;87(1):15-9. Epub 2006/01/13.
6. Pandyan AD, Johnson GR, Price CI, Curless RH, Barnes MP, Rodgers H. A review of the properties and limitations of the Ashworth and modified Ashworth Scales as measures of spasticity. *Clinical rehabilitation*. 1999;13(5):373-83. Epub 1999/09/25.
7. Rizzo MA, Hadjimichael OC, Preiningerova J, Vollmer TL. Prevalence and treatment of spasticity reported by multiple sclerosis patients. *Mult Scler*. 2004;10(5):589-95. Epub 2004/10/09.
8. Guy W, Cleary PA. Pretreatment status and its relationship to the length of drying-out period. *Psychopharmacology bulletin*. 1976;12(2):20-2. Epub 1976/04/01.
9. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain*. 2004;8(4):283-91. Epub 2004/06/23.
10. Schwid SR, Goodman AD, Apatoff BR, Coyle PK, Jacobs LD, Krupp LB, et al. Are quantitative functional measures more sensitive to worsening MS than traditional measures? *Neurology*. 2000;55(12):1901-3. Epub 2001/01/03.
11. EuroQolGroup. Health Questionnaire - English version for the UK (validated for Ireland). 1990.

12. 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. *Journal of neurology, neurosurgery, and psychiatry*. 2005;76(1):58-63. Epub 2004/12/21.
13. Vickrey BG, Hays RD, Genovese BJ, Myers LW, Ellison GW. Comparison of a generic to disease-targeted health-related quality-of-life measures for multiple sclerosis. *Journal of clinical epidemiology*. 1997;50(5):557-69. Epub 1997/05/01.
14. Vickrey BG, Hays RD, Harooni R, Myers LW, Ellison GW. A health-related quality of life measure for multiple sclerosis. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*. 1995;4(3):187-206. Epub 1995/06/01.
15. Attarian H. Importance of sleep in the quality of life of multiple sclerosis patients: a long under-recognized issue. *Sleep medicine*. 2009;10(1):7-8. Epub 2008/05/17.
16. Lobentanz IS, Asenbaum S, Vass K, Sauter C, Klosch G, Kollegger H, et al. Factors influencing quality of life in multiple sclerosis patients: disability, depressive mood, fatigue and sleep quality. *Acta neurologica Scandinavica*. 2004;110(1):6-13. Epub 2004/06/08.
17. Brass SD, Duquette P, Proulx-Therrien J, Auerbach S. Sleep disorders in patients with multiple sclerosis. *Sleep medicine reviews*. 2010;14(2):121-9. Epub 2009/11/03.
18. Nicholl L, Hobart J, Dunwoody L, Cramp F, Lowe-Strong A. Measuring disability in multiple sclerosis: is the Community Dependency Index an improvement on the Barthel Index? *Mult Scler*. 2004;10(4):447-50. Epub 2004/08/26.
19. Hsieh YW, Wang CH, Wu SC, Chen PC, Sheu CF, Hsieh CL. Establishing the minimal clinically important difference of the Barthel Index in stroke patients. *Neurorehabilitation and neural repair*. 2007;21(3):233-8. Epub 2007/03/14.
20. Pirpiris M, Wilkinson AJ, Rodda J, Nguyen TC, Baker RJ, Nattrass GR, et al. Walking speed in children and young adults with neuromuscular disease: comparison between two assessment methods. *Journal of pediatric orthopedics*. 2003;23(3):302-7. Epub 2003/05/02.
21. Paltamaa J, Sarasoja T, Leskinen E, Wikstrom J, Malkia E. Measures of physical functioning predict self-reported performance in self-care, mobility, and domestic life in ambulatory persons with multiple sclerosis. *Archives of physical medicine and rehabilitation*. 2007;88(12):1649-57. Epub 2007/12/01.
22. Musselman KE, Fouad K, Misiaszek JE, Yang JF. Training of walking skills overground and on the treadmill: case series on individuals with incomplete spinal cord injury. *Physical therapy*. 2009;89(6):601-11. Epub 2009/05/09.
23. Gartlehner G, Moore CG. Direct versus indirect comparisons: a summary of the evidence. *Int J Technol Assess Health Care*. 2008;24(2):170-7. Epub 2008/04/11.
24. Sutton A, Ades AE, Cooper N, Abrams K. Use of indirect and mixed treatment comparisons for technology assessment. *Pharmacoeconomics*. 2008;26(9):753-67. Epub 2008/09/05.
25. Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ*. 2009;338:b1147. Epub 2009/04/07.
26. Salanti G, Marinho V, Higgins JPT. A case study of multiple-treatments meta-analysis demonstrates that covariates should be considered. *J Clin Epidemiol*. 2009;62(8):857-64. Epub 2009/01/23.
27. Pharmaceuticals G. NCT00711646: A Study of Sativex® for Relief of Spasticity in Subjects With Multiple Sclerosis. 2002.

28. Pharmaceuticals G. EudraCT: 2004-002509-63: A double blind, randomised, placebo controlled, parallel group study of Sativex, in subjects with symptoms of spasticity due to multiple sclerosis. 2005; Available from: <https://www.clinicaltrialsregister.eu/ctr-search/>.
29. Pharmaceuticals G. NCT00681538: A Study of the Safety and Effectiveness of Sativex®, for the Relief of Symptoms of Spasticity in Subjects With Multiple Sclerosis (MS). 2008.
30. Sapienza UoRL. NCT00202423: Use of Cannabinoids in Patients With Multiple Sclerosis. 2005 [updated 28. November 2005]; Available from: <http://clinicaltrials.gov/ct2/show/NCT00202423?term=NCT00202423&rank=1>.
31. Wade DT, Makela P, Robson P, House H, Bateman C. Do cannabis-based medicinal extracts have general or specific effects on symptoms in multiple sclerosis? A double-blind, randomized, placebo-controlled study on 160 patients. *Mult Scler*. 2004;10(4):434-41. Epub 2004/08/26.
32. Pharmaceuticals G. Clinical Study Report GWMS0106: A double blind, randomised, parallel group study to assess the efficacy, safety and tolerability of Cannabis Based Medicine 1:1 THC:CBD compared with placebo for the treatment of spasticity in patients with multiple sclerosis. 2006 NCT00711646.
33. Collin C, Davies P, Mutiboko IK, Ratcliffe S. Randomized controlled trial of cannabis-based medicine in spasticity caused by multiple sclerosis. *European journal of neurology : the official journal of the European Federation of Neurological Societies*. 2007;14(3):290-6. Epub 2007/03/16.
34. Pharmaceuticals G. Clinical Study Report GWCL0403: A double blind, randomised, placebo controlled, parallel group study of Sativex in subjects with symptoms of spasticity due to multiple sclerosis. 2006 EudraCT Nummer: 2004-002509-63.
35. Collin C, Ehler E, Waberszinek G, Alsindi Z, Davies P, Powell K, et al. A double-blind, randomized, placebo-controlled, parallel-group study of Sativex, in subjects with symptoms of spasticity due to multiple sclerosis. *Neurological research*. 2010;32(5):451-9. Epub 2010/03/24.
36. Pharmaceuticals G. Clinical Study Report GWSP0604: A Two-phase, Phase 3 Study of the Safety and Efficacy of Sativex®, in the Symptomatic Relief of Spasticity in Subjects with Spasticity due to Multiple Sclerosis: Phase A - Single-blind Response Assessment; Phase B - Double-blind, Randomised, Placebo Controlled, Parallel Group Study. 2009.
37. Novotna A, Mares J, Ratcliffe S, Novakova I, Vachova M, Zapletalova O, et al. A randomized, double-blind, placebo-controlled, parallel-group, enriched-design study of nabiximols* (Sativex®), as add-on therapy, in subjects with refractory spasticity caused by multiple sclerosis. *European journal of neurology : the official journal of the European Federation of Neurological Societies*. 2011;18(9):1122-31. Epub 2011/03/03.
38. Aragona M, Onesti E, Tomassini V, Conte A, Gupta S, Gilio F, et al. Psychopathological and cognitive effects of therapeutic cannabinoids in multiple sclerosis: a double-blind, placebo controlled, crossover study. *Clinical neuropharmacology*. 2009;32(1):41-7. Epub 2008/11/04.
39. Briefing Document: Sativex Oromucosal Spray and International Law. 2011.
40. Novartis Pharma: Fachinformation Lioresal® Intrathecal 0,05 mg/1 ml Injektionslösung. Stand: November 2011.
41. Dario A, Tomei G. A benefit-risk assessment of baclofen in severe spinal spasticity. *Drug safety : an international journal of medical toxicology and drug experience*. 2004;27(11):799-818. Epub 2004/09/08.

42. Pharmaceuticals G. NCT00702468: Evaluate the Maintenance of Effect After Long-term Treatment with Sativex® in Subjects With Symptoms of Spasticity Due to Multiple Sclerosis 2008; Available from:

<http://clinicaltrials.gov/ct2/show/NCT00681538?term=NCT00681538&rank=1>.

43. Notcutt W, Langford R, Davies P, Ratcliffe S, Potts R. A placebo-controlled, parallel-group, randomised withdrawal study of subjects with symptoms of spasticity due to multiple sclerosis who are receiving long-term Sativex(R) (nabiximols). *Mult Scler*. 2011. Epub 2011/09/01.

44. Pharmaceuticals G. Clinical Study Report GWSP0702: A placebo controlled, parallel group, randomised withdrawal study of subjects with symptoms of spasticity due to multiple sclerosis who are receiving long-term GW-1000-02 (Sativex®). 2009.

45. Pharmaceuticals G. Clinical Study Report GWN10005: An "N of 1" placebo controlled study of cannabis based medicine extracts in treatment of multiple sclerosis, spinal cord injury, peripheral nerve injury or rheumatoid arthritis. 2003.

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 ¹⁷] – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	Meglitinide/	848
2	Nateglinide/	1686
3	Repaglinide/	2118
4	(glinid* or meglitinid* or nateglinid* or repaglinid*).ab,ti.	1069
5	(starlix or novonorm or novo norm or prandin).ab,ti.	32
6	(105816-04-4 or 135062-02-1).rn.	2854
7	or/1-6	3467
8	Diabetes mellitus/	224164
9	Non Insulin dependent Diabetes mellitus/	91081
10	(diabet* or niddm or t2dm).ab,ti.	379777
11	or/8-10	454517
12	(random* or double-blind*).tw.	650136
13	placebo*.mp.	243550
14	or/12-13	773621
15	and/7,11,14	719

¹⁷ Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455.

Literaturrecherche für: Randomisierte Studien mit THC/CBD zur Behandlung von Spastik bei MS (4.3.1.1 und 4.3.2.3.1)		
Datenbankname	MEDLINE	
Suchoberfläche	PUBMED	
Datum der Suche	03.11.2011	
Zeitsegment	bis 2011 Woche 44	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ¹⁷]- Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	Sativex	54
2	nabiximols	8
3	cannabinoids	9823
4	OR/ 1-3	9847
5	baclofen	5693
6	tizanidine	361
7	dantrolene	2155
8	tolperisone	95
9	tetrazepam	83
10	OR/5-9	8167
11	OR/4,10	17980
12	multiple sclerosis	44225
13	AND/11,12	485
14	randomized controlled trial [Publication Type]	296801
15	randomized [Title/Abstract]	238642
16	OR/14,15	393265
17	AND/13,16	65

Literaturrecherche für Randomisierte Studien mit THC/CBD zur Behandlung von Spastik bei MS (4.3.1.1 und 4.3.2.3.1)		
Datenbankname	EMBASE	
Suchoberfläche	Elsevier	
Datum der Suche	01.11.1011	
Zeitsegment	bis 2011 Woche 44	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ¹⁷]- Strategy minimizing difference between sensitivity and specificity Filter für EMBASE/lim	
#	Suchbegriffe	Ergebnis
1	Sativex	16652
2	nabiximols	17
3	cannabinoids	31091
4	OR/ 1-3	31092
5	baclofen	11505
6	tizanidine	1716
7	dantrolene	4662
8	tolperisone	371
9	tetrazepam	541
10	OR/5-9	16789
11	OR/4,10	47530
12	multiple sclerosis	1319
13	AND/11,12	13
14	random:ab:ti	117246
15	double blind:ab:ti	107418
16	randomized controlled trial/exp	267000
17	OR/14-16	421130
18	AND/13,17	1

Literaturrecherche für: Randomisierte Studien mit THC/CBD zur Behandlung von Spastik bei MS (4.3.1.1 und 4.3.2.3.1)		
Datenbankname	Cochrane	
Suchoberfläche	Cochrane Library	
Datum der Suche	01.11.2011	
Zeitsegment	bis 2011 Woche 44	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ¹⁷]- Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	Sativex	26
2	nabiximols	3
3	cannabinoids	167
4	OR/1-3	183
5	baclofen	326
6	tizinadine	117
7	dantrolene	83
8	tolperisone	20
9	tetrazepam	36
10	OR/5-9	510
11	OR/4,10	685
12	MeSH descriptor multiple sclerosis	1506
13	AND/11,12	67
14	(randomized controlled trial):pt	302184
15	MeSH descriptor Randomized Controlled Trial as Topic	15390
16	OR/14,15	312320
17	AND/13, 16	40

Anhang 4-B: Suchstrategien – Suche in Studienregistern

Geben Sie nachfolgend die Suchstrategien für die in Abschnitt 4.2.3.3 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

Die Suchstrategien wurden für die Abschnitte 4.3.1.1 und 4.3.2.3.1 durchgeführt.

Studienregisterrecherche in Clinicaltrials	
Studienregister	Clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov/ct2/search/advanced
Datum der Suche	04.10.2011
Suchstrategie	(Sativex* OR GW*1000* OR Cannab* OR *THC* Nabiximol* OR Marinol* OR tetrahydrocannabinol* OR Nabilon*) [ALL-FIELDS] AND (multiple sclerosis OR spasticity OR Spasm*) [CONDITIONS] AND (All Studies) [Recruitment] AND (All Studies) [Study results] AND (All Studies) [Study Type]
Treffer	16
Feld	Suchbegriff
Search Terms	Sativex* OR GW*1000* OR Cannab* OR *THC* OR Nabiximol* OR Marinol* OR tetrahydrocannabinol* OR Nabilon*
Recruitment	All Studies
Study results	All Studies
Study type	All Studies
Conditions	multiple sclerosis OR spasticity OR spasm*

Studienregisterrecherche in Clinicalstudyresults	
Studienregister	Clinicalstudyresults.org
Internetadresse	http://www.clinicalstudyresults.org
Datum der Suche	04.10.2011
Suchstrategie	(muscle spasm) [Studied Indications or Disease]
Treffer	1
Feld	Suchbegriff
Studied Indications or Disease	Muscle spasm

Studienregisterrecherche in Clinicalstudyresults	
Studienregister	Clinicalstudyresults.org
Internetadresse	http://www.clinicalstudyresults.org
Datum der Suche	04.10.2011
Suchstrategie	(Multiple Sclerosis) [Studied Indications or Disease]
Treffer	12
Feld	Suchbegriff
Studied Indications or Disease	Multiple Sclerosis

Studienregisterrecherche in International Clinical Trials Registry Platform Search Portal	
Studienregister	International Clinical Trials Registry Platform Search Portal
Internetadresse	http://apps.who.int/trialsearch/AdvSearch.aspx
Datum der Suche	04.10.2011
Suchstrategie	(multiple sclerosis OR spasticity OR spasm*) [Condition] AND (Sativex* OR GW*1000* OR Cannab* OR *THC* OR Nabiximol* OR Marinol* OR tetrahydrocannabinol* OR Nabilon*) [Title] AND (ALL) [Recruitment Status]
Treffer	20 records for 19 trials found
Feld	Suchbegriff
Title	Sativex* OR GW*1000* OR Cannab* OR *THC* OR Nabiximol* OR Marinol* OR tetrahydrocannabinol* OR Nabilon*
Condition	multiple sclerosis OR spasticity OR spasm*
Recruitment status	ALL

Studienregisterrecherche in EU Clinical Trials Register	
Studienregister	EU Clinical Trials Register
Internetadresse	https://www.clinicaltrialsregister.eu/ctr-search/
Datum der Suche	04.10.2011
Suchstrategie	(multiple sclerosis OR spasticity OR spasm*) AND (Sativex* OR GW*1000* OR Cannab* OR *THC* OR Nabiximol* OR Marinol* OR tetrahydrocannabinol* OR Nabilon*)
Treffer	92
Feld	Suchbegriff
Search for CTA	(multiple sclerosis OR spasticity OR spasm*) AND (Sativex* OR GW*1000* OR Cannab* OR *THC* OR Nabiximol* OR Marinol* OR tetrahydrocannabinol* OR Nabilon*)

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.

Publikationstitel	Ausschlussgrund
Killestein J, Hoogervorst EL, Reif M, Kalkers NF, Van Loenen AC, Staats PG, Gorter RW, Uitdehaag BM, Polman CH (2002): Safety, tolerability, and efficacy of orally administered cannabinoids in MS.	Die Studie präsentiert Daten zur Behandlung mit der Wirkstoffkombination THC:CBD, die nicht im Verhältnis 1:1 zur Behandlung eingesetzt wurde.
Zajicek JP, Fox P, Sanders H, Wright D, Vickery D, Nunn A, Thompson A (2003): Cannabinoids for treatment of spasticity and other symptoms related to multiple sclerosis (CAMS study): multicenter randomized placebo-controlled trial.	Die Studie präsentiert Daten zur Behandlung mit der Wirkstoffkombination THC:CBD, die nicht im Verhältnis 1:1 zur Behandlung eingesetzt wurde.
Zajicek JP, Sanders HP, Wright DE, Vickery PJ, Ingram WM, Reilly SM, Nunn AJ, Teare LJ, Fox PJ, Thompson AJ (2005): Cannabinoids in multiple sclerosis (CAMS) study: safety and efficacy data for 12 months follow up.	Die Studie präsentiert Daten zur Behandlung mit der Wirkstoffkombination THC:CBD, die nicht im Verhältnis 1:1 zur Behandlung eingesetzt wurde.
Wade DT; Makela PM; House H; Bateman C; Robson P (2006). Long-term use of a cannabis-based medicine in the treatment of spasticity and other symptoms in multiple sclerosis.	Es handelt es sich um eine nicht randomisierte Studie.
Notcutt W, Langford R, Davies P, Ratcliffe S, Potts R (2011): A placebo-controlled, parallel-group, randomised withdrawal study of subjects with symptoms of spasticity due to multiple sclerosis who are receiving long-term Sativex® (nabiximols)	Diese Studie betrachtete keinen patientenrelevanten Endpunkt. Es handelt sich um eine Absetzungsstudie.

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 4-89 (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])	Register-eintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
GWN 10005	Nein	Ja	Nein	Ja(45)	Nein	Nein
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
Es liegt keine aktiv kontrollierte abgebrochene Studie vor.						
<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>						

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 4-90 (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						
placebokontrolliert						
NEU-CAN-04 fMRI and Neurophysiological Study Protocol on Cannabinoids in Multiple Sclerosis	nein	nein	Ja [S. Andrea Hospital]	nein	Ja (30)	nein
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
Es liegt keine laufende aktiv kontrollierte RCT-Studie vor.						
<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>						

Diese Studie wurde nach Suche in den Studienregistern identifiziert (Tabelle 4-4).

Anhang 4-F: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie.

Tabelle 4-91 (Anhang): Studiendesign und -methodik für Studie GWMS0001

Item a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	<p>Primäre Zielsetzung: Vergleich der Wirksamkeit von GW-1000-02 (Sativex® synonym: Sativex® Spray zur Anwendung in der Mundhöhle) gegenüber Placebo in der Linderung der fünf Schlüsselsymptome nach 6 Wochen Therapie bei MS-Patienten anhand der visuellen Analogskala (VAS), die die Schwere der Symptome misst.</p> <p>Sekundäre Zielsetzungen: Vergleich der Wirksamkeit von GW-1000-02 gegenüber Placebo in der Linderung der fünf Symptome der MS nach 6 Wochen Therapie mit Hilfe der VAS Bestimmung der Verträglichkeit von GW-1000-02 in MS-Patienten basierend auf unerwünschten Ereignissen</p>
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Multizentrische, doppelblinde, randomisierte, Placebo-kontrollierte, parallelgruppierte Studie. Patienten wurden im Verhältnis 1:1 zu GW-1000-02 oder Placebo randomisiert.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p>Nach Studienbeginn, aber vor Randomisierung, wurde die Zusammensetzung der sonstigen Bestandteile von GW-1000-02 und Placebo geändert.</p> <p>Es wurden weitere Studienzentren an der Studie beteiligt, die Packungsbeilage und die Beschriftung der Medikation geändert und zwar nachdem 43 der 160 Patienten randomisiert wurden.</p> <p>Es wurden nachträglich Patienten von der Studie ausgeschlossen, die entweder Fentanyl oder Sildenafil einnahmen, da der Metabolismus dieser Medikamente durch das Cytochrome P450-System vermittelt wird, welches lt. Tierversuchsstudien bei hohen Dosen von CBD inhibiert wird.</p>
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Inhaltliche relevante Einschlusskriterien:</p> <ul style="list-style-type: none"> • Männlich oder weiblich, über 18 Jahre alt. • MS jeden Typs. • Stabiler Krankheitsverlauf über mindestens 4 Wochen vor Studienstart. • Symptome mit bestimmter Schwere (>50 mm auf einer 100mm VAS Schwere-skala) in mindestens einer der spezifischen Kategorien: Spastizität, Muskelspasmen, gestörter Blasenkontrolle, neuropathischen Schmerzen, Tremor der Extremitäten. • Stabiles Behandlungsregime der anti-spastischen Medikation über mindestens vier Wochen vor Studienbeginn und die Bereitschaft, diese im Studienverlauf fortzusetzen. • Kein Konsum von Cannabinoiden (Cannabis, Marinol® oder Nabilone) mindestens 7 Tage vor Studienbeginn. • Bereitschaft, effektive Verhütung während der Studie und den drei folgenden Monaten zu benutzen; oder es erfolgte eine chirurgische Sterilisation oder weibliche Patienten in der Postmenopause. • klinisch akzeptable Laborresultate beim Vorstudien-Screening. • Bereitwillig und in der Lage sein, alle Studienbedingungen zu erfüllen und durchzuführen. • Bereitwillig und in der Lage sein, die betreffenden Informationen, die

Item a	Charakteristikum	Studieninformation
		<p>Einverständniserklärung zu lesen, in Erwägung zu ziehen und zu verstehen und anschließend die schriftliche Einverständniserklärung abzugeben. Probanden, die nicht in der Lage waren zu lesen oder die Verfahrensanweisungen zu unterschreiben, wurden wie in der Erklärung von Helsinki beschrieben, behandelt.</p> <ul style="list-style-type: none"> • Bereitwilligkeit, den Allgemeinarzt und gegebenenfalls den Facharzt über die Teilnahme an der Studie zu informieren. • Bereitwilligkeit, dass der Name dem Home-Office anlässlich der Teilnahme an der Studie mitgeteilt wird. <p>Inhaltlich relevante Ausschlusskriterien:</p> <ul style="list-style-type: none"> • offensichtlicher oder vermuteter Drogenmissbrauch, einschließlich Alkohol. • momentane oder vergangene Abhängigkeit von Cannabis. • bekannte oder vermutete Nebenwirkungen bei Cannabinoiden, die Psychosen oder andere schwere psychische Störungen hervorrufen. • Historie irgendeiner Art von Schizophrenie, anderen psychotischen Erkrankungen oder bedeutsamen psychiatrischen Erkrankungen oder Persönlichkeitsstörungen ohne Depressionen assoziiert mit einer chronischen Krankheit. • Einnahme von Medikamenten, die Levopoda enthalten (Sinemeth[®], Sinemet plus[®], Levopoda[®], L-dopa[®], Madopar[®], Benserazide[®]). • schwere kardiovaskuläre Funktionsstörungen einschließlich Angina, unkontrollierte Hypertonie oder eine unkontrollierte symptomatische Herzrhythmusstörung. • erhebliche Nieren- oder Leberstörungen laut medizinischer Historie bzw. laut Laborergebnissen. • Vorerkrankung an Epilepsie. • Endstadium der Krankheit oder andere Bedingungen, bei der eine Medikation mit einem Placebo unangebracht wäre. • Schwangerschaft, Stillzeit oder Risiko einer Schwangerschaft. • Teilnahme in einer anderen klinischen Studie während der 12 Wochen vor Studieneintritt. • geplanter Krankenhausaufenthalt während dem Studieneintritt und Visite 6. • geplante Reise außerhalb der UK während des Studieneintritts und Visite 6.

Item a	Charakteristikum	Studieninformation
4b	Studienorganisation und Ort der Studiendurchführung	Sponsor: GW Pharma Ltd. Ort der Studiendurchführung: UK
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p><u>Gruppe: GW-1000-02 (Sativex®):</u> Bestandteile: THC (27 mg/ml)/CBD (25 mg/ml) als Extrakt von <i>Cannabis sativa</i> mit Pfefferminzöl , 0.05% (v/v), in Ethanol:Propylenglycol (50:50). Applikation in Hüben, jeder Hub setzt 100 µl (THC 2.7 mg und CBD 2.5 mg).</p> <p><u>Placebo</u> Bestandteile: Pfefferminzöl , 0.05% (v/v), Chinolingelb, 0.005% (w/v), sunset yellow, 0.0025% (w/v), in Ethanol:Propylenglycol (50:50). Applikation in Hüben, jeder Hub setzt 100 µl Flüssigkeit frei.</p> <p>Die Studienmedikation war ein Pumpspray zur Anwendung in der Mundhöhle. Die maximal erlaubte Dosis waren 8 Hübe über 3 Stunden und 48 Hübe über 24 Stunden. Die optimale Dosis war den Studienteilnehmern überlassen in Abhängigkeit von Wirksamkeit und Verträglichkeit und maximal erlaubter Dosis. Die Studienmedikation wurde von den Studienteilnehmern selbst appliziert, bzw. wo notwendig, durch die betreuende Person appliziert.</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 Endpunkt Die primäre Messung der Wirksamkeit war die Änderung der Intensität der individuell identifizierten primären Beeinträchtigung von der Basislinie bis Woche 6 (Ende des doppelblinden Abschnitts). Der Basislinienwert war der Durchschnitt aller Werte der Visiten 1 und 2 und der Woche 6-Wert der Visite 4.</p> <p>Sekundäre Endpunkte Jeder der fünf individuellen Arten der Beeinträchtigung wurde analysiert und wie für die primäre Beeinträchtigung präsentiert.</p> <p>Andere sekundäre Endpunkte:</p> <ol style="list-style-type: none"> 1. Bewertung der Beeinträchtigung anhand von Tagebüchern mit Hilfe des 100mm VAS 2. Vergleiche innerhalb der Gruppe als deskriptive Statistik 3. Assessment Battery Data 4. Assessment Battery Vergleich zwischen den Gruppen zur Visite 4: <ol style="list-style-type: none"> 1. Guy's Neurological Disability Scale (GNDS) zur Messung der klinischen Körperbehinderung der MS-Patienten 2. Barthel Activities of Daily Living (ADL) Index zur Messung der Selbstbetreuung und Mobilitätsfähigkeit 3. Aitkinson Morley Information Processing Battery (AMIPB) zur Bestimmung der Kognition und mentalen Aufmerksamkeit 4. Beck Depression Inventory (BDI) zur Messung der zwei primären Dimensionen der Depression 5. Fatigue Severity Scale (FSS) zur Messung des Erschöpfungslevels aufgrund von neurologischen Krankheiten

Item a	Charakteristikum	Studieninformation
		<ol style="list-style-type: none"> 6. 28-item General Health Questionnaire (GHQ-28) zur Feststellung von nicht-psychotischen mentalen Erkrankungen (Ängstlichkeit und Depression) 7. Fragenkatalog zur Blasenkontrolle 8. Schlaf-Fragenkatalog zur Bestimmung der Qualität und Quantität des Schlafes mit Hilfe der VAS 5. Assessment Battery Vergleich innerhalb der Gruppe zu den Visiten 4 und 5 <ol style="list-style-type: none"> 1. Caregiver Strain Index(CSI) zur Bestimmung der Strapazen der betreuenden Personen 2. Nine-Hole Peg Test zur Bestimmung der Geschicklichkeit und Koordination 3. Rivermead Mobility Index (RMA) als Maß für die Selbstmobilisierung der Patienten 4. Schlaf-Fragenkatalog zur Bestimmung der Qualität und Quantität des Schlafes mit Hilfe der VAS 6. Patient Global Opinion of Effect on MS (PGIC) für MS zur Beschreibung von „viel besser“, „besser“, „keine Veränderung“, „schlecht“, „viel schlechter“ 7. Modified Ashworth Scale zum Test der Gelenkflexibilität durch einen Arzt und zur Bestimmung des Grads der Spastizität in spezifischen Muskelgruppen 8. Therapeutische Spanne der Dosen 9. Sicherheit:Fragenkatalog zur Bestimmung der Nebenwirkungen und schweren Nebenwirkungen 10. Klinische Labordaten 11. Vitalparameter 12. Grad der Berauschtigkeit 13. EKG
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Es wurde nach Studienbeginn keine Änderung der Zielkriterien vorgenommen.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahl dieser Studie basiert auf dem Vergleich der beiden randomisierten Behandlungen der primären Beeinträchtigung. Es wurde die Änderung des VAS von der Basislinie für die primäre Beeinträchtigung benutzt und die Standardabweichung von 24mm angenommen. Basierend auf diesen Annahmen, waren 55 Patienten pro Behandlungsarm notwendig, um eine Differenz in der Änderung des VAS von der Basislinie von 15mm mit einer Power von 90% bei einem zweiseitigen Signifikanztest mit $\alpha=0,05$ ermitteln zu können. Eine Ausfallrate von ungefähr 15% wurde erwartet. Es wurde berechnet, dass 130 Patienten rekrutiert werden sollten.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	NA

Item a	Charakteristikum	Studieninformation
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Zur Durchführung der Randomisierung wurden randomisiert permutierte Blöcke der Größe 4 generiert und wurden durch die primäre Beeinträchtigung und das Klinikzentrum stratifiziert.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Patienten wurden 1:1 entweder zur Behandlungsgruppe GW-1000-02 (Sativex®) oder zu Placebo randomisiert.
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	Dem Apotheker jedes Zentrums wurde ein Randomisierungsschema für jede primäre Beeinträchtigung und der festgesetzten Behandlung in Form der sequentiellen Abfolge der Patientennummern der verwendeten Randomisierungsliste bereitgestellt.
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 Gruppenzuteilung erfolgte in 4er Blöcken. Die Patienten wurden in jedem Zentrum sequentiell den randomisierten Patientennummern zugeteilt. Es ist nicht bekannt, wer die Zuteilung durchführte.
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?	Während der 6wöchigen Behandlungsphase waren die Patienten/Probanden, diejenigen, die die Behandlung durchführten und die Zielgrößen beurteilten, verblindet. Die Studienmedikation wurde mit dem Namen des Sponsors beschriftet und enthielt den Studiencode, die Patientenummer, die Nummer der Visite und das Verfallsdatum. Die Patienten- und Visitenummer wurde durch den Apotheker angebracht, der freien Zugang zum Randomisierungsplan hatte. Der Untersuchende und die klinische Abteilung (GW Pharma) hatten für jeden Patienten verschlossene Umschläge mit Codierungsbruch.
11b	Falls relevant, Beschreibung der	Placebo und Verum wurden beide als Spray mit gleicher Färbung verabreicht.

Item a	Charakteristikum	Studieninformation
	Ähnlichkeit von Interventionen	
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Die primäre Analyse aller Wirksamkeitsendpunkte (primärer und sekundärer) wurde mit der intention-to-treat (ITT) Population durchgeführt. Imputationen für missing values wurden nicht vorgenommen.</p> <p>Alle statistischen Gruppenvergleiche wurden auf dem 5% Signifikanzniveau durchgeführt. Multiples Testen wurde nicht durchgeführt.</p> <p>Die Analysen wurden mit Methoden der deskriptiven Statistik durchgeführt (n, Mittelwert, SD, Median, Minimum und Maximum oder absolute und relative Häufigkeit).</p> <p>Die Änderungen in der Beurteilung der Intensität der primären Beeinträchtigung auf der VAS wurden in einem ANCOVA Modell untersucht. Das Modell enthielt Behandlungsgruppe und Zentrum als Faktoren sowie den Baseline VAS-Wert als Covariate.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	NA
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe Flow-Chart unten
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>a) Gruppe 1: 80 Gruppe 2: 80 b) Gruppe 1: 80 Gruppe 2: 80 c) Gruppe 1: 77 Gruppe 2: 77</p>
13b	Für jede Gruppe: Beschreibung von	Siehe Flow-Chart unten

Item a	Charakteristikum	Studieninformation
	verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	05/2001 – 07/2002 Run-in Phase: 1-2 Wochen Im Anschluss 6-wöchige aktive Behandlungsphase Im weiteren Anschluss 4 Wochenopen-label Phase Danach Möglichkeit der Teilnahme an einer Langzeitstudie
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde wie geplant beendet.
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

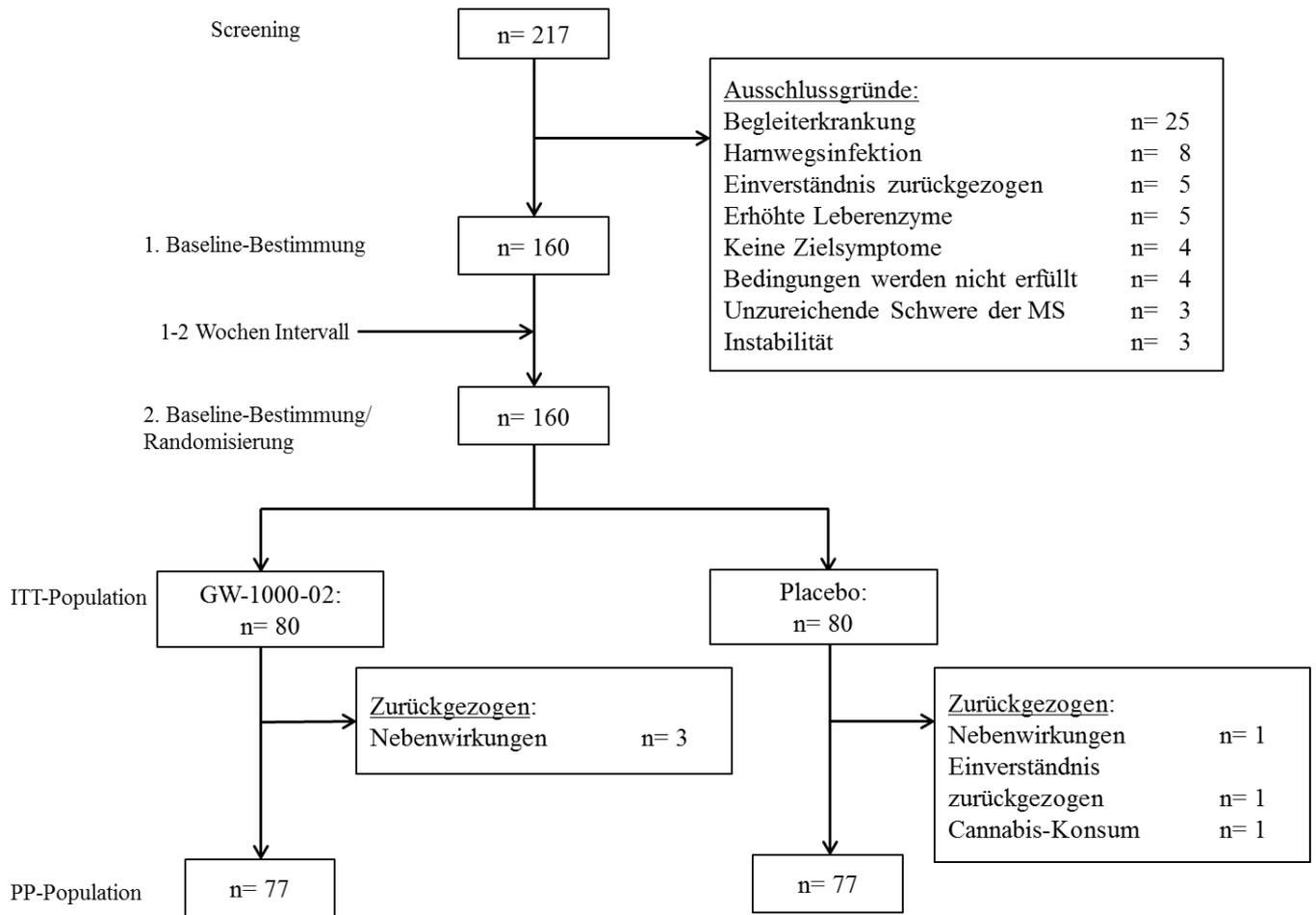


Tabelle 4-92 (Anhang): Studiendesign und -methodik für Studie GWMS0106

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	<p>Primäre Zielsetzung: Vergleich der Wirksamkeit von GW-1000-02 (Sativex® synonym: Sativex® Spray zur Anwendung in der Mundhöhle) gegenüber Placebo auf den Grad der Muskelspastizität basierend auf dem täglichen Beurteilungswert auf der 11-stufigen Numerischen Bewertungsskala (NRS).</p> <p>Sekundäre Zielsetzungen: Die Evaluierung der Wirkung von GW-1000-02 im Vergleich zu Placebo auf:</p> <ul style="list-style-type: none"> • Ashworth Skala. • Die tägliche Anzahl von Spastiken auf einer 5-stufigen ordinalen Skala. • Motricity Index. • Die Globalbeurteilung der Veränderung durch den Patienten (SGIC) <p>Die Beurteilung der Sicherheit und Verträglichkeit von GW-1000-02</p> <p>Es wurden keine statistischen Hypothesen getestet</p>
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Multizentrische, doppelt-blinde, randomisierte, Placebo-kontrollierte, parallelgruppierte Studie. Patienten wurden im Verhältnis 2:1 zu GW-1000-02 oder Placebo randomisiert.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Folgende Änderungen wurden vor Beginn der Studie implementiert: Die primäre Endpunktmessung (ursprünglich Visuelle Analogskala) wurde ersetzt durch die Änderung des durchschnittlichen Spastikscores gemessen auf der 11-stufigen NRS. Die Ashworth Skala wurde als sekundärer Endpunkt erhoben.
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Inhaltliche relevante Einschlusskriterien:</p> <ul style="list-style-type: none"> • Männlich oder weiblich, über 18 Jahre alt • Stabiler Krankheitsverlauf über mindestens 3 Monate vor Studienstart • Bestehende Multiple Sklerose mit nicht zufriedenstellend eingestellter Spastik. • Signifikante Spastik in mindestens 2 Muskelgruppen, definiert als Wert von zwei oder mehr auf der Ashworth-Skala für jede Muskelgruppe. • Stabiles Behandlungsregime der anti-spastischen Medikation über mindestens 30 Tage vor Studienbeginn und die Bereitschaft, diese im Studienverlauf fortzusetzen • Kein Konsum von Cannabinoiden (Cannabis, Marinol® oder Nabilone) mindestens 7 Tage vor Studienbeginn. • Keine Krankheitsgeschichte von Schizophrenie, andern psychotischen Erkrankungen, schweren Persönlichkeitsstörungen oder anderen signifikanten psychischen Erkrankungen außer Depression. • klinisch akzeptable Laborwerte zur Visite 2. • Kein Alkohol- oder Drogenmissbrauch. • Keine schweren Herz-Kreislauf-Erkrankungen. • Keine Krankheitsgeschichte von Epilepsie.

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none">• Keine signifikante Nieren- oder Leberinsuffizienz.• Keine regelmäßige Einnahme von Levodopa (Sinemet[®], Sinemet Plus[®], Levodopa, L-dopa, Madopar[®], Benserazide) mindestens 7 Tage vor Studienteilnahme.• Keine Einnahme von Sildenafil (Viagra[®]) während der Studie.• Keine Einnahme von Fentanyl (Durogesic[®], Actiq[®])• Keine Einnahme von Antiarrhythmika

Item ^a	Charakteristikum	Studieninformation
4b	Studienorganisation und Ort der Studiendurchführung	Sponsor: GW Pharma Ltd.; 8 Zentren in UK und 4 Zentren Rumänien
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p><u>Gruppe: GW-1000-02 (Sativex®):</u> Bestandteile: THC (27 mg/ml)/CBD (25 mg/ml) als Extrakt von <i>Cannabis sativa</i> L., mit Pfefferminzöl , 0.05% (v/v), in Ethanol:Propylenglycol (50:50). Applikation in Hüben, jeder Hub setzt 100 µl (THC 2.7 mg und CBD 2.5 mg).</p> <p><u>Placebo</u> Bestandteile: Pfefferminzöl , 0.05% (v/v), Chinolingelb, 0.005% (w/v), sunset yellow, 0.0025% (w/v), in Ethanol:Propylenglycol (50:50). Applikation in Hüben, jeder Hub setzt 100 µl Flüssigkeit frei.</p> <p>Die Studienmedikation war ein Pumpspray zur Anwendung in der Mundhöhle. Die maximal erlaubte Dosis waren 8 Hübe über 3 Stunden und 48 Hübe über 24 Stunden. Die optimale Dosis war den Studienteilnehmern überlassen in Abhängigkeit von Wirksamkeit und Verträglichkeit und maximal erlaubter Dosis. Die Studienmedikation wurde von den Studienteilnehmern selbst appliziert, bzw. wo notwendig, durch die betreuende Person appliziert.</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 Endpunkt Die primäre Messung der Wirksamkeit erfolgte auf täglicher Beurteilung des Grads der Spastik auf einer validierten 11-stufigen NRS im Patiententagebuch.</p> <p>Sekundäre Endpunkte Ein kombinierter Score aus Ashworth Skala in den von Spastik betroffenen Muskelgruppen an Visit 1, Visit 2, Visit 3 und Visit 4. Die tägliche Häufigkeit von Spastiken dokumentiert im Tagebuch auf einer ordinalen Skala von 0-4. Globale Beurteilung der klinischen Veränderung durch den Patienten am Studienende. Analyse der Responder (Anteil der Patienten mit 30 % und 50 % Verbesserung der NRS Bewertung der Spastik seit Studienstart). Zeit bis zur Verbesserung der NRS Beurteilung Safety Variablen Die Verträglichkeit und Sicherheit der Behandlung wurde anhand folgender Parameter beurteilt: Auftreten von Nebenwirkungen, Blut- und Urinanalyse, Vitalzeichen, EKG, körperliche Untersuchungsbefunde <i>Zu den Erhebungszeitpunkten siehe Flow Chart unten</i></p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Das primäre Zielkriterium wurde vor Studienstart geändert. Basierend auf neuen Erkenntnissen bestand die Möglichkeit, dass das ursprüngliche Instrument (Ashworth Skala) zu insensitive ist, um kleine aber klinisch signifikante Effekt auf die Spastik zu detektieren. Als Resultat der Änderung des primären Zielkriteriums wurde die Fallzahl um 45 auf 174 Patienten erhöht. Der ursprüngliche Plan für eine Zwischenanalyse des Ashworth Scores wurde entsprechend verworfen.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahl wurde auf Basis der primären Variable (Änderung der Beurteilung des Grads der Spastik auf der 11-stufigen NRS) kalkuliert. Es wurde eine SD von 2,1 für die Änderung auf der 11-stufigen NRS

Item ^a	Charakteristikum	Studieninformation
		angenommen. Die Randomisierung wurde im Verhältnis 2:1 durchgeführt, um mehr Informationen über die Verum Studienmedikation zu gewinnen. Bei Anwendung eines zwei-seitigen Tests auf dem 5 % Niveau und einer SD von 2,1 war eine Fallzahl von 159 (106 Verum, 53 Placebo) Patienten ausreichend gewesen, um mit 80 % Power eine Differenz zwischen den Behandlungsgruppen von 1,0 für die 11-stufige NRS Skala eine Bewertung der Spastizität zu erkennen.

Item ^a	Charakteristikum	Studieninformation
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	NA
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Zur Durchführung der Randomisierung wurde ein vorgefertigter Randomisierungscode verwendet, der durch den Statistiker von GW erzeugt wurde. Die Gruppenteilung erfolgte in 3er Blöcken.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Patienten wurden 2:1 entweder zur Behandlungsgruppe GW-1000-02 (Sativex®) oder zu Placebo randomisiert.
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 Studienmedikation war vorkonfektioniert durch das GW Clinical Trials Supplies Department und gekennzeichnet mit der Patientennummer. Die Patienten wurden in jedem Zentrum sequenziell den randomisierten Patientennummern zugeteilt.
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?	Der Randomisierungscode wurde durch einen Statistiker von GW erzeugt. Die Gruppenteilung erfolgte in 3er Blöcken. Die Patienten wurden in jedem Zentrum sequenziell den randomisierten Patientennummern 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,	Patienten und Studienärzte/-personal waren verblindet (doppelt-blindes Studiendesign) Verblindung: Die Studienmedikation wurde in 5,5 ml Typ I Braunglasfläschchen bereitgestellt, gekennzeichnet mit dem Namen von GW, dem Studiencode, der Patientennummer, Visitennummer, Referenznummer und Verfallsdatum.

Item ^a	Charakteristikum	Studieninformation
	wie wurde die Verblindung vorgenommen?	
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Placebo und Verum wurden beide als Spray mit gleicher Färbung verabreicht.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Die primäre Analyse aller Wirksamkeitsendpunkte (primärer und sekundärer) wurde mit der intention-to-treat (ITT) Population durchgeführt. Imputationen für missing values wurden nicht vorgenommen.</p> <p>Alle statistischen Gruppenvergleiche wurden auf dem 5% Signifikanzniveau durchgeführt. Multiples Testen wurde nicht durchgeführt.</p> <p>Die Analysen wurden mit Methoden der deskriptiven Statistik durchgeführt (n, Mittelwert, SD, Median, Minimum und Maximum oder absolute und relative Häufigkeit).</p> <p>Die Änderungen in der Spastizitätsbeurteilung auf der 11-stufigen NRS Skala wurden in einem ANCOVA Modell untersucht. Das Modell enthielt Behandlungsgruppe und Zentrum als Faktoren sowie den Baseline NRS Wert als Covariate.</p> <p>Der durchschnittliche Score für die Anzahl Spastiken am Tag wurde im gleichen ANCOVA Modell ausgewertet wie der primäre Endpunkt NRS.</p> <p>Für die Globalbeurteilung der klinischen Veränderung durch den Patienten (SGIC) wurde der Anteil Patienten mit Beurteilung 'Very Much Improved', 'Much Improved' or 'Minimally Improved' zwischen den Gruppen durch den Fisher's Exact Test verglichen.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	NA
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe Flow-Chart unten
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert	<p>a) Gruppe 1: 124 Gruppe 2: 65</p> <p>b) Gruppe 1: 124 Gruppe 2: 65</p> <p>c) Gruppe 1: 124 Gruppe 2: 65</p>

Item ^a	Charakteristikum	Studieninformation
	wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flow-Chart unten
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	06/2002 – 03/2004 Run-in Phase: 2 Wochen Im Anschluss 6-wöchige aktive Behandlungsphase Im weiteren Anschluss 7-10 Tage Nachbeobachtung
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde wie geplant beendet.
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

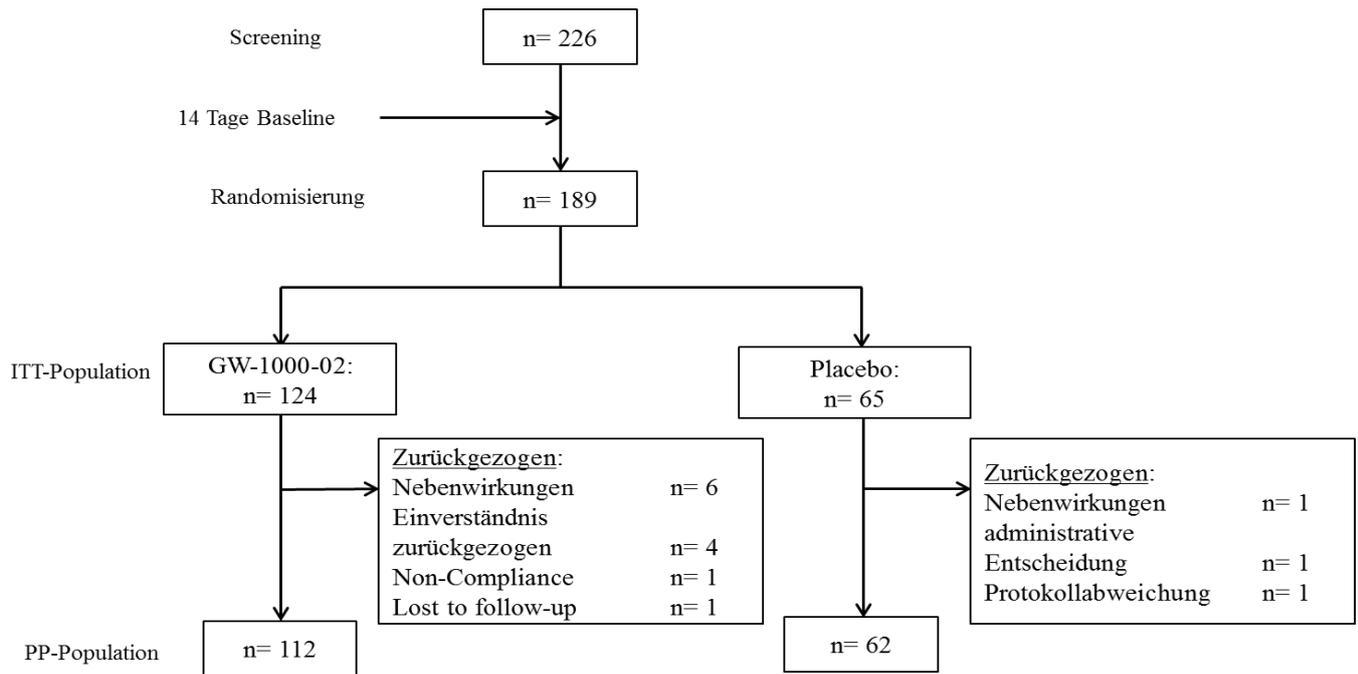


Tabelle 4-93 (Anhang): Studiendesign und -methodik für Studie GWCL0403

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	<p>Primäre Zielsetzung: Vergleich der Wirksamkeit von GW-1000-02 (Sativex® synonym: Sativex® Spray zur Anwendung in der Mundhöhle) gegenüber Placebo auf die Verbesserung der Spastizität bei Patienten mit MS.</p> <p>Sekundäre Zielsetzungen: Die Evaluierung der Wirkung von GW-1000-02 im Vergleich zu Placebo auf:</p> <ol style="list-style-type: none"> 1. Sekundäre Parameter der Spastizität. 2. Anzahl der Patienten mit mindestens 30 % oder mindestens 50 % Verbesserung des primären Endpunkts 3. Funktionale Parameter 4. Lebensqualität <p>Die Beurteilung der Sicherheit und Verträglichkeit von GW-1000-02</p> <p>Die getestete Nullhypothese war, dass kein Behandlungsgruppenunterschied besteht in der durchschnittlichen NRS Beurteilung der Spastik. Bei einem p-Wert von unter 0,05 wurde statistische Signifikanz der Gruppenunterschiede angenommen und die Nullhypothese wurde verworfen und die Alternativhypothese angenommen.</p>
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnissen	Multizentrische, doppelt-blinde, randomisierte, Placebo-kontrollierte, parallelgruppierte Studie. Patienten wurden im Verhältnis 1:1 zu GW-1000-02 oder Placebo randomisiert.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Keine
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Inhaltliche relevante Einschlusskriterien:</p> <ul style="list-style-type: none"> • Männlich oder weiblich, über 18 Jahre alt • Stabiler Krankheitsverlauf über mindestens 3 Monate vor Studienstart • Bestehende Multiple Sklerose seit mindestens 6 Monaten • Bestehende nicht zufriedenstellend eingestellte Spastik seit mindestens 3 Monaten. • Stabiles Behandlungsregime der anti-spastischen Medikation über mindestens 30 Tage vor Studienbeginn und die Bereitschaft, diese im Studienverlauf fortzusetzen. • Stabiles Behandlungsregime mit Disease Modifying Drugs über mindestens 6 Monate und die Bereitschaft, diese im Studienverlauf fortzusetzen. • Keine Begleiterkrankungen mit Symptomen von Spastik • Keine Konsum von Cannabis mindestens 30 Tage vor Studienbeginn. • Keine Konsum von Cannabinoid Medikamenten mindestens 60 Tage vor

Item ^a	Charakteristikum	Studieninformation
		<p data-bbox="568 280 730 309">Studienbeginn.</p> <ul data-bbox="568 309 1406 611" style="list-style-type: none"><li data-bbox="568 309 1406 398">• Keine Krankheitsgeschichte von Schizophrenie, anderen psychotischen Erkrankungen, schweren Persönlichkeitsstörungen oder anderen signifikanten psychischen Erkrankungen außer Depression.<li data-bbox="568 398 986 427">• Kein Alkohol- oder Drogenmissbrauch.<li data-bbox="568 427 1406 495">• Keine schweren Herz-Kreislauf-Erkrankungen, insbesondere Herzinfarkt in den vergangenen 12 Monaten.<li data-bbox="568 495 1023 524">• Keine Krankheitsgeschichte von Epilepsie.<li data-bbox="568 524 1102 553">• Keine signifikante Nieren- oder Leberinsuffizienz.<li data-bbox="568 553 1406 611">• Keine Behandlung mit Botulinum Toxin und keine Bereitschaft, diese Behandlung während der Studie zu unterbrechen.

Item ^a	Charakteristikum	Studieninformation
4b	Studienorganisation und Ort der Studiendurchführung	Sponsor: GW Pharma Ltd.; 14 Zentren in UK und 8 Zentren in Tschechien
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<u>Gruppe: GW-1000-02 (Sativex®):</u> Bestandteile: THC (27 mg/ml)/CBD (25 mg/ml) als Extrakt von <i>Cannabis sativa</i> L., mit Pfefferminzöl, 0.05% (v/v), in Ethanol:Propylenglycol (50:50). Applikation in Hüben, jeder Hub setzt 100 µl (THC 2.7 mg und CBD 2.5 mg). <u>Placebo:</u> Bestandteile: Pfefferminzöl, 0.05% (v/v), Chinolingelb, 0.005% (w/v), sunset yellow, 0.0025% (w/v), in Ethanol:Propylenglycol (50:50). Applikation in Hüben, jeder Hub setzt 100 µl Flüssigkeit frei.
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</p> <p>Die primäre Messung der Wirksamkeit erfolgte auf täglicher Beurteilung des Grads der Spastik auf einer 11-stufigen NRS im Patiententagebuch. Der primäre Endpunkt „Grad der Spastik“ wurde auf der 11-stufigen NRS Skala im Patiententagebuch dokumentiert und als Durchschnitt der letzten 14 Tage ausgewertet.</p> <p>Die sekundären Wirksamkeitsmessungen waren:</p> <p>Modifizierte Ashworth-Skala, 10m-Gehstreckentest, Barthel-Index (Alltagsfunktionen, ADL), EDSS CGIC Schlafqualität Lebensqualität (EQ-5D, MSQol-54 [nur UK]) Nebenwirkungen</p> <p>Visitenplan:</p> <p>Run-in: 1 Woche 14-wöchige aktive Behandlungsphase: 28 Tage nach post follow-up visit</p> <ol style="list-style-type: none"> 1. Visit: Baseline 2. Visit: 7 Tage 3. Visit: nach 14 Tagen Behandlung 4. Visit: nach 6 Wochen Behandlung 5. Visit: nach 10 Wochen Behandlung 6. Visit: nach 14 Wochen Behandlung 7. Visit: 28 Tage nach Visit 6 oder Abbruch der Studie <p>Die regelmäßige Monitorierung der Studienzentren erfolgte durch GW Clinical Department. Clinical Research Associates (CRAs) waren verantwortlich für die Überprüfung der Protokolladhärenz, Einhaltung von ICH GCP und Datenvollständigkeit und -richtigkeit. Die original CRFs und andere essentielle Dokumente wurden zu GW gesendet.</p> <p>Vor Studienbeginn wurde ein Investigator Meeting durch GW durchgeführt. Ein Training erfolgte hinsichtlich Studienprotokoll, Studienprozeduren,</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Arzneimittelsicherheitsberichterung, ICH GCP, EU Directives CRF completion und Studienmedikationsbelieferung.</p> <p>Die Laboruntersuchungen wurden in einem Zentrallabor durchgeführt.</p> <p>Die CRF Daten wurden in einer elektronischen Datenbank übertragen. Manuelle und elektronische Prüfungen der Daten wurden durchgeführt, die Datenvalidierung wurde in einem Datenvalidierungsplan spezifiziert und dokumentiert. Änderungen wurden per Audit Trail dokumentiert.</p> <p>Unabhängige Audits wurden seitens GW Clinical Quality Assurance Department in sechs der Studienzentren durchgeführt. Weiter wurden interne Systemaudits bezüglich Sicherheitsberichterung, Monitoring, Studienlogistik der statistischen Analyse sowie der Datenbank durchgeführt.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	<i>Nicht zutreffend</i>
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahl wurde auf Basis der primären Variable (Änderung der Beurteilung des Grads der Spastik auf der 11-stufigen NRS) kalkuliert. Es wurde von einem NRS Wert von 6.0 an Baseline ausgegangen. Eine 30 % Abnahme der NRS wurde als klinisch relevant angesehen, entsprechend einer Reduktion um 1,8 Punkte in dieser Studie. Wenn 25 % der Patienten eine solche Response zeigen und die übrigen 75 % hätten keine Veränderung, würde dass einer durchschnittlichen Abnahme von 0,45 Punkten entsprechen. Basierend auf den bisherigen Studienergebnissen sowie Angaben aus der Literatur wird von einer Gruppendifferenz von mindestens 0,75 bei einer SD von 2,0 ausgegangen. Unter diesen Annahmen waren auf dem 5 % Irrtumsniveau bei einer Power von 80 % eine Fallzahl von 226 Patienten (113 in jeder Gruppe) erforderlich, um diesen Unterschied von 0,75 zu erkennen. Unter der Annahme von 20 % ungeeigneten Patienten müssten 284 Patienten (142 in jeder Gruppe) randomisiert werden.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	NA
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Randomisierungssequenz wurde vor Studienbeginn computer-basiert durch das GW Biometrics Department erzeugt (Software: SAS Version 9.1).
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierung wurde 1:1 per Blockrandomisierung durchgeführt und für jedes Zentrum zusätzlich stratifiziert nach dem Kriterium Ambulanz/Nicht-Ambulanz. Der jeweils erste Block war vom Umfang 2, die jeweils nachfolgenden vom Umfang 4 je Zentrum (Randomisierungsschema für maximal 48 Patienten je Zentrum).
9	Randomisierung, Geheimhaltung der Behandlungsfolge	Die Studienmedikation war vorkonfektioniert durch GW Clinical Trials Supplies Department und gekennzeichnet mit der Patientenummer. Die Patienten wurden in jedem Zentrum sequenziell den randomisierten

Item ^a	Charakteristikum	Studieninformation
	(allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Patientennummern zugeteilt.
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 Gruppenzuteilung zu den vorgefertigten Patientencodes erfolgte durch das Personal in jedem Zentrum.
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) Probanden waren verblindet. b) Investigatoren waren verblindet. c) Probanden und Investigatoren beurteilten die Zielgrößen, sie waren verblindet. Verblindung: Die Studienmedikation wurde in 5,5 ml Typ I Braunglasfläschchen bereitgestellt, gekennzeichnet mit dem Namen von GW, dem Studiencode, der Patientennummer, Visitennummer, Referenznummer und Verfallsdatum.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Placebo und Verum wurden beide als Spray mit gleicher Färbung verabreicht.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die primäre Analyse aller Wirksamkeitsendpunkte (primärer, sekundäre und tertiäre) wurden mit der intention-to-treat (ITT) Population durchgeführt. Alle statistischen Gruppenvergleiche wurden auf dem 5% Signifikanzniveau durchgeführt. Multiple Vergleiche wurden nicht durchgeführt. Die Endpunktanalyse erfolgte per last observation carried forward (LOCF). Die Analysen wurden mit Methoden der deskriptiven Statistik durchgeführt (n, Mittelwert, SD, Median, Minimum und Maximum oder absolute und relative Häufigkeit).

Item ^a	Charakteristikum	Studieninformation
		<p>Die Änderungen in der Spastizitätsbeurteilung (Durchschnitt der letzten 6 Tage (Baseline) bzw. der letzten 7-14 Tage (Endpunkt)) auf der 11-stufigen NRS Skala wurden in einem ANCOVA Modell untersucht. Für Patienten ohne Post-Baseline Wert wurde der Baseline-Wert eingesetzt. Das Modell enthielt Behandlungsgruppe, Zentrum und Ambulanzenstatus (ja/nein) und Vorgeschichte von Cannabiskonsum als Hauptfaktoren sowie den Baseline NRS Wert als Covariate. Das initiale Modell wurde auf mögliche, zu berücksichtigende Interaktionen zwischen den Hauptfaktoren geprüft; irrelevante Interaktionen wurden auf dem 10 % Niveau eliminiert.</p> <p>Für die Anzahl der Responder wurden als Vergleichsgröße Gruppendifferenzen und Odds ratio inklusive 95 %-CI dargestellt.</p> <p>Die Schlafqualität wurde als Gruppendifferenz der Veränderung mit dem dazu gehörigen 95% CI dargestellt. Dieser Endpunkt wurde in einem ANCOVA Modell untersucht. Das Modell enthielt Behandlungsgruppe, Zentrum und Ambulanzenstatus (ja/nein) als Hauptfaktoren sowie den Baseline NRS Wert als Covariate.</p> <p>Der 10 m-Gehstreckentest und der Barthel Index wurden mit einem ANCOVA Modell untersucht. Das Modell enthielt Behandlungsgruppe, Zentrum und Ambulanzenstatus (ja/nein) als Hauptfaktoren sowie den Baseline NRS Wert als Covariate.</p> <p>Für die Globalbeurteilung der klinischen Veränderung durch die betreuende Person (auf der 7-stufigen CGIC Skala) wurden die zwei Behandlungsgruppen mittels ordinaler logistischer Regression verglichen. Das Modell enthielt Behandlungsgruppe, Zentrum und Ambulanzenstatus (ja/nein) Faktoren. Odds ratio sowie 95 %-CI und p-Wert wurden präsentiert.</p> <p>Der EQ-5D VAS Score wurde als Gruppendifferenz der Veränderung mit dem dazugehörigen 95 %-CI dargestellt. Dieser Endpunkt wurde in einem ANCOVA Modell untersucht. Das Modell enthielt Behandlungsgruppe, Zentrum und Ambulanzenstatus (ja/nein) als Hauptfaktoren sowie den Baseline NRS Wert als Covariate.</p>

Item ^a	Charakteristikum	Studieninformation
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Siehe 12a
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe Flow-Chart unten
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) Gruppe 1: 167 Gruppe 2: 170 b) Gruppe 1: 167 Gruppe 2: 170 c) Gruppe 1: 166 Gruppe 2: 169
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flow-Chart unten
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	03/2005 - 12/2005
14b	Informationen, warum die Studie endete oder beendet	Studie wurde planmäßig beendet.

Item ^a	Charakteristikum	Studieninformation
	wurde	
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

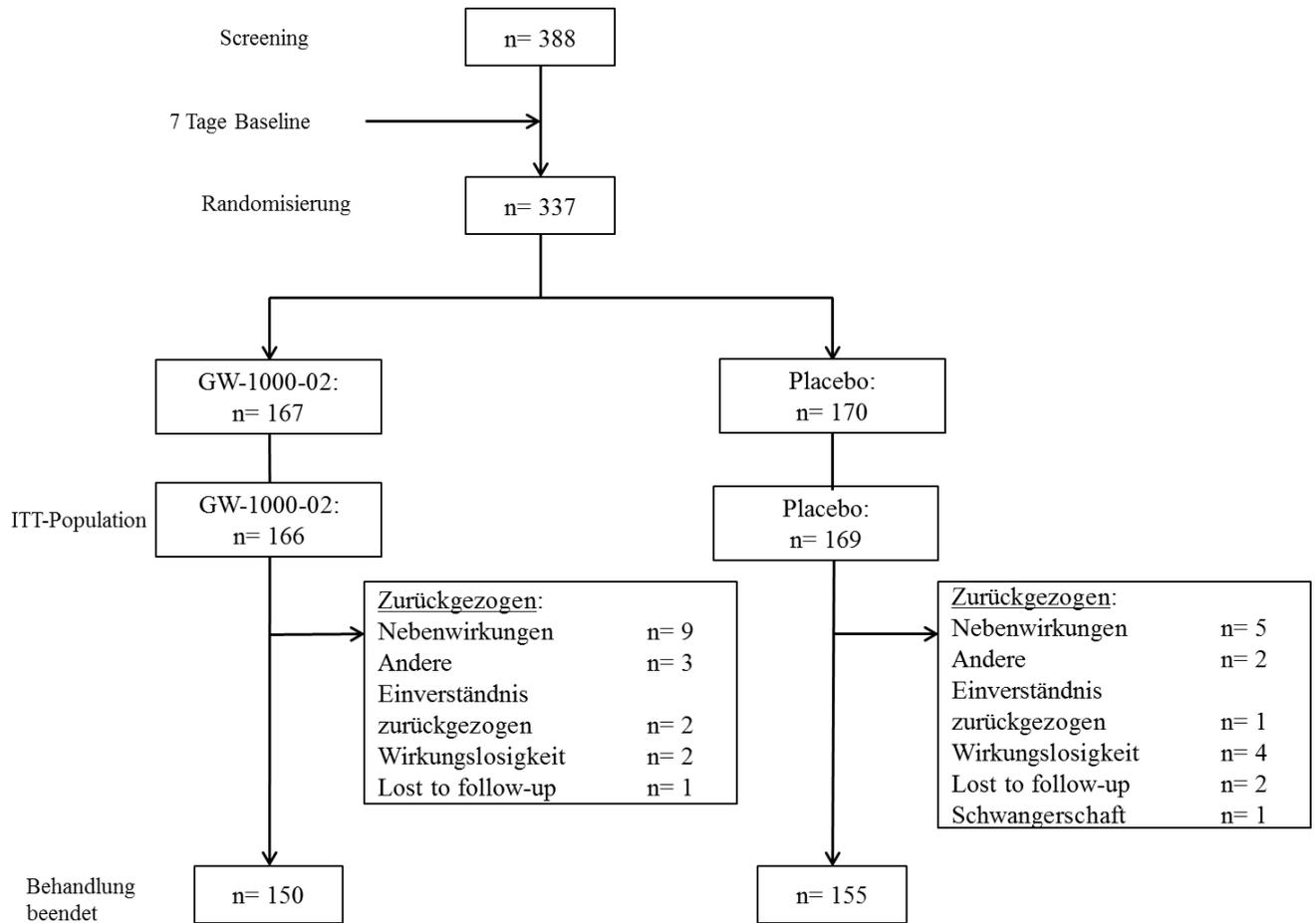


Tabelle 4-94 (Anhang): Studiendesign und -methodik für Studie GWSP0604

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	<p>Primäre Zielsetzung: Vergleich der Wirksamkeit von GW-1000-02 (Sativex[®] synonym: Sativex[®] Spray zur Anwendung in der Mundhöhle) gegenüber Placebo auf die Verbesserung der Spastizität bei Patienten mit MS, die auf GW-1000-02 positiv ansprechen (Responder).</p> <p>Sekundäre Zielsetzungen: Die Evaluierung der Wirkung von GW-1000-02 im Vergleich zu Placebo auf: 5. Sekundäre Parameter der Spastizität. 6. Funktionale Parameter der Spastizität.</p> <p>Die Beurteilung der Sicherheit und Verträglichkeit von GW-1000-02</p> <p>In der Phase A (Verum Run-in Phase) wurde keine statistische Hypothese getestet. In der Phase B (randomisierte Parallelgruppenphase) wurde die Nullhypothese getestet, dass kein Behandlungsgruppenunterschied besteht in der durchschnittlichen NRS Beurteilung der Spastik. Bei einem p-Wert von unter 0,05 wurde statistische Signifikanz der Gruppenunterschiede angenommen und die Nullhypothese wurde verworfen und die Alternativhypothese angenommen.</p>
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	<p>Multizentrische, doppelt-blinde, randomisierte, Placebo-kontrollierte, parallelgruppierte Studie mit einer einarmigen, einfach verblindeten Verum Run-in Phase.</p> <p>Die Patienten wurden im Verhältnis 1:1 zu GW-1000-02 oder Placebo randomisiert.</p> <p>Run-in Phase A diente zur Identifizierung von Respondern auf GW-1000-02 und war einarmig, einfach verblindet über 4 Wochen.</p> <p>Phase B war doppelt-blind, Placebo-kontrolliert, 1:1 randomisiert in Parallelgruppen über 12 Wochen in Respondern aus Phase A.</p>
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p>Das initiale Studienprotokoll wurde dahingehend geändert, dass als Baselinewerte für die Auswertung nicht mehr die Werte zu Beginn der Phase A genutzt wurden, sondern die Werte zu Beginn der Phase B (Ende Phase A). Dies geschah, damit die Analysen besser auf die Testhypothese zutrafen und der Baselinewert näher am Zeitpunkt der Randomisierung lag zur Präzision der Schätzung in der Analyse. Entsprechend musste die Fallzahl angepasst werden.</p> <p>Das ursprüngliche Protokoll ermöglichte nur den Einschluss von ambulanten Patienten. Dies wurde geändert, wonach auch nicht ambulante Patienten eingeschlossen werden konnten, wodurch das zukünftig angenommene klinische Einsatzgebiet von GW-1000-02 besser wiedergegeben werden sollte.</p>
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p>Inhaltlich relevante Einschlusskriterien:</p> <ul style="list-style-type: none"> • Männlich oder weiblich, über 18 Jahre alt • Stabiler Krankheitsverlauf über mindestens 3 Monate vor Studienstart • Bestehende Multiple Sklerose seit mindestens 6 Monaten • Moderate Spastik von mindestens 4 Punkten auf der 11-stufigen NRS für

Item ^a	Charakteristikum	Studieninformation
		<p>den Schweregrad der Spastik</p> <ul style="list-style-type: none"> • Bestehende nicht vollständig eingestellte Spastik seit mindestens 3 Monaten. • Behandelt mit anti-spastischer Medikation oder Therapieversuch mit anti-spastischer Medikation hat versagt oder wurde nicht vertragen • Stabiles Behandlungsregime der anti-spastischen Medikation über mindestens 30 Tage vor Studienbeginn und die Bereitschaft, diese im Studienverlauf fortzusetzen • Stabiles Behandlungsregime mit Disease Modifying Drugs über mindestens 3 Monate und die Bereitschaft, diese im Studienverlauf fortzusetzen • Patient hatte in der Verum Run-in Phase mindestens 20 % Rückgang des NRS Scores der Spastizität gezeigt, um randomisiert werden zu können. • Patient musste in der Run-in Phase vollständig compliant gewesen sein mit der geforderten Dokumentation des Schweregrads der Spastik auf der NRS • Keine Begleiterkrankungen mit Symptomen von Spastik • Kein Konsum von Cannabis oder Cannabinoid Medikamenten oder Acomplia[®] mindestens 30 Tage vor Studienbeginn. • Keine Krankheitsgeschichte von Schizophrenie, anderen psychotischen Erkrankungen, schweren Persönlichkeitsstörungen oder anderen signifikanten psychischen Erkrankungen außer Depression. • Kein Alkohol- oder Drogenmissbrauch. • Keine schweren Herz-Kreislauf-Erkrankungen, insbesondere Herzinfarkt oder Herzfunktionsstörungen in den vergangenen 12 Monaten. • Keine Krankheitsgeschichte von nicht gut eingestellter Epilepsie oder wiederkehrenden Krämpfen. • Keine signifikante Nieren- oder Leberinsuffizienz. • Keine Behandlung mit Botulinum Toxin.

Item ^a	Charakteristikum	Studieninformation
4b	Studienorganisation und Ort der Studiendurchführung	Sponsor: GW Pharma Ltd.; 52 Studienzentren (18 Zentren in UK, 11 Zentren in Spanien, 10 Zentren in Polen, 8 Zentren in Tschechien, 5 Zentren in Italien)
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Phase A: GW-1000-02 (Sativex[®]): Bestandteile: THC (27 mg/ml)/CBD (25 mg/ml) als Extrakt von <i>Cannabis sativa</i> L., mit Pfefferminzöl , 0,05% (v/v), in Ethanol:Propylenglycol (50:50). Applikation in Hüben, jeder Hub setzt 100 µl (THC 2.7 mg und CBD 2.5 mg) frei.</p> <p>Phase B:</p> <p><u>Gruppe: GW-1000-02 (Sativex[®]):</u> Bestandteile: THC (27 mg/ml)/CBD (25 mg/ml) als Extrakt von <i>Cannabis sativa</i> L., mit Pfefferminzöl , 0,05% (v/v), in Ethanol:Propyleneglycol (50:50). Applikation in Hüben, jeder Hub setzt 100 µl (THC 2.7 mg und CBD 2.5 mg) frei.</p> <p><u>Placebo:</u> Bestandteile: Pfefferminzöl , 0,05% (v/v), Chinolingelb, 0,005% (w/v), sunset yellow, 0,0025% (w/v), in Ethanol:Propylenglycol (50:50). Applikation in Hüben, jeder Hub setzt 100 µl Flüssigkeit frei.</p> <p>Die Studienmedikation sollte individuell selbst dosiert werden mit einer maximalen Anzahl Sprühstöße von 12 pro Tag.</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 Endpunkt</p> <p>Die primäre Messung der Wirksamkeit erfolgte als tägliche Beurteilung des Grads der Spastik auf der 11-stufigen NRS im Patiententagebuch. Der primäre Endpunkt Grad der Spastik wurde auf der 11-stufigen NRS Skala im Patiententagebuch dokumentiert und als Durchschnitt der letzten 7 Tage ausgewertet.</p> <p>Die sekundären Wirksamkeitsmessungen waren:</p> <p>Anteil der Patienten mit Verbesserung des Grads der Spastik um mindestens 30 % und mindestens 50 % auf der 11-stufigen NRS</p> <p>Anzahl der Spastiken pro Tag</p> <p>Beurteilung der Schlafbeeinträchtigung (11-stufige NRS)</p> <p>10m-Gehstreckentest,</p> <p>Barthel-Index (Alltagsfunktionen, ADL),</p> <p>Beurteilung der globalen Veränderung durch Patienten, betreuende Person und Ärzte (SGIC, CGIC, PhGIC).</p> <p>Tertiäre Parameter:</p> <p>Lebensqualität EQ-5D und SF-36</p> <p>Nebenwirkungen</p> <p>Die regelmäßige Monitorierung der Studienzentren erfolgte durch GW Clinical Department. Clinical Research Associates (CRAs) waren verantwortlich für die Überprüfung der Protokolladhärenz, Einhaltung von ICH GCP und Datenvollständigkeit und –richtigkeit. Die original CRFs und andere essentielle Dokumente wurden zu GW gesendet.</p> <p>Vor Studienbeginn wurden Investigator Meetings durch GW durchgeführt. Ein Training erfolgte hinsichtlich Studienprotokoll, Studienprozeduren, Arzneimittelsicherheitsberichterung, ICH GCP und EU Directives CRF.</p> <p>Die Laboruntersuchungen wurden in einem Zentrallabor durchgeführt.</p> <p>Die CRF Daten wurden in eine elektronische Datenbank übertragen.</p>

Item ^a	Charakteristikum	Studieninformation
		Manuelle und elektronische Prüfungen der Daten wurden durchgeführt, die Datenvalidierung wurde in einem Datenvalidierungsplan spezifiziert und dokumentiert. Änderungen wurden per Audit Trail dokumentiert. Unabhängige Audits wurden seitens GW Clinical Quality Assurance Department durchgeführt in 13 der Studienzentren. Weiter wurden regelmäßige Supplier Audits durchgeführt bezüglich Datenmanagement, Statistischer Auswertung, Zentrallabor, Sicherheitsberichterstattung, Monitoring, Studienlogistik der statistischen Analyse sowie der Datenbank, IVRS Provider. Weiter wurden interne Systemaudits durchgeführt bezüglich Sicherheitsberichterstattung, statistischer Auswertung, Studienlogistik, regulatorischer Belange, sowie Dokumentenkontrolle und Monitoring Systeme.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nein
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahl wurde auf Basis der primären Variable (Änderung der Beurteilung des Grads der Spastik auf der 11-stufigen NRS) kalkuliert. Basierend auf den bisherigen Studienergebnissen wurde von einer Gruppendifferenz von mindestens 0,75 bei einer SD von 1,6 ausgegangen. Unter diesen Annahmen war auf dem 5 % Niveau bei einer Power von 90 % eine Fallzahl von 194 Patienten (97 in jeder Gruppe) erforderlich um in Phase B diesen Unterschied von 0,75 zu erkennen. Unter der Annahme von 20 % ungeeigneter Patienten mussten 244 Patienten (122 in jeder Gruppe) randomisiert werden. Zusätzlich wurde von einer Responderrate von 50 % in der Verum Run-in Phase A ausgegangen, so dass von 488 Patienten ausgegangen wurde, die in Phase A in die Studie aufgenommen werden mussten.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	NA
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Randomisierungssequenz wurde vor Studienbeginn computer-basiert durch einen unabhängigen Statistiker erzeugt.
8b	Einzelheiten (z. B. Blockrandomisierung , Stratifizierung)	Die Randomisierung wurde 1:1 als Blockrandomisierung durchgeführt, balanciert und stratifiziert nach Zentren.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte	Die Studienmedikation war vorkonfektioniert und gekennzeichnet mit der Patientennummer. Die Nummern wurden den Patienten gemäß IVRS zugeteilt.

Item ^a	Charakteristikum	Studieninformation
	Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur 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?	Zuteilung durch IVRS.
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) Phase A: Patienten b) Phase B: Patienten und Behandler c) keine weiteren Personen außer Patienten und Behandler Verblindung: Die Studienmedikation wurde in 5.5 ml Typ I Braunglasfläschchen bereitgestellt, gekennzeichnet mit dem Namen von GW, dem Studiencode, der Patientenummer, Visitennummer, Referenznummer und Verfallsdatum.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Placebo und Verum wurden beide als Spray mit gleicher Färbung verabreicht.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Ergebnisse aus der Run-in-Phase A wurden mittels deskriptiver Methoden statistisch ausgewertet. Die Auswertung der Phase B erfolgte mittels der im Folgenden beschriebenen statistischen Methoden. Die Analyse aller Wirksamkeitsendpunkte (primärer, sekundäre und tertiäre) wurden mit der intention-to-treat (ITT) Population durchgeführt. Alle statistischen Gruppenvergleiche wurden auf dem 5 % Signifikanzniveau durchgeführt. Multiple Vergleiche wurden zum Testen der Hypothese für den primären Endpunkt nicht durchgeführt. Alle Endpunkte wurden mit Methoden der deskriptiven Statistik dargestellt (n, Mittelwert, SD, Median, Minimum und Maximum oder absolute und relative Häufigkeit).

Item ^a	Charakteristikum	Studieninformation
		<p>Der primäre Wirksamkeitsendpunkt, Änderung in der Spastizitätsbeurteilung (Durchschnitt der letzten 7 Tage) auf der 11-stufigen NRS Skala, wurde in einem ANCOVA Modell getestet. Das Modell enthielt Behandlungsgruppe, Land und Ambulanzenstatus (ja/nein) als Hauptfaktoren sowie den Baseline Wert als Covariate.</p> <p>Für die Anzahl Responder (30 % und 50 % Verbesserung der NRS Beurteilung der Spastik) wurden Odds ratios inklusive 95 %-CI dargestellt.</p> <p>Die Anzahl Spastiken pro Tag wurde mit demselben ANCOVA Modell analysiert wie der primäre Endpunkt.</p> <p>Die Schlafqualität wurde als Gruppendifferenz der Veränderung mit dem dazu gehörigen 95% CI dargestellt und mit demselben ANCOVA Modell analysiert wie der primäre Endpunkt.</p> <p>Der 10 m-Gehstreckentest und der Barthel Index wurden mit einem ANCOVA Modell untersucht. Das Modell enthielt Behandlungsgruppe, Zentrum und Ambulanzenstatus (ja/nein) als Hauptfaktoren sowie den Baseline NRS Wert als Covariate.</p> <p>Für die Globalbeurteilung der klinischen Veränderung durch Patient, betreuende Person und Arzt (auf jeweils 7-stufiger Skala) wurden die zwei Behandlungsgruppen mittels ordinaler logistischer Regression verglichen.</p> <p>EQ-5D VAS und SF-36 Scores wurde als Gruppendifferenz der Veränderung mit dem dazugehörigen 95 %-CI dargestellt und mit demselben ANCOVA Modell analysiert wie der primäre Endpunkt.</p>

Item ^a	Charakteristikum	Studieninformation
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	NA
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Siehe Flow Chart
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	Phase B: (Responder aus Phase A) a) Gruppe 1: 124 Gruppe 2: 117 b) Gruppe 1: 124 Gruppe 2: 117 c) Gruppe 1: 124 Gruppe 2: 117
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flow-Chart unten
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	01/2008 – 01/2009
14b	Informationen, warum die Studie endete oder beendet wurde	Studie wurde planmäßig beendet.
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

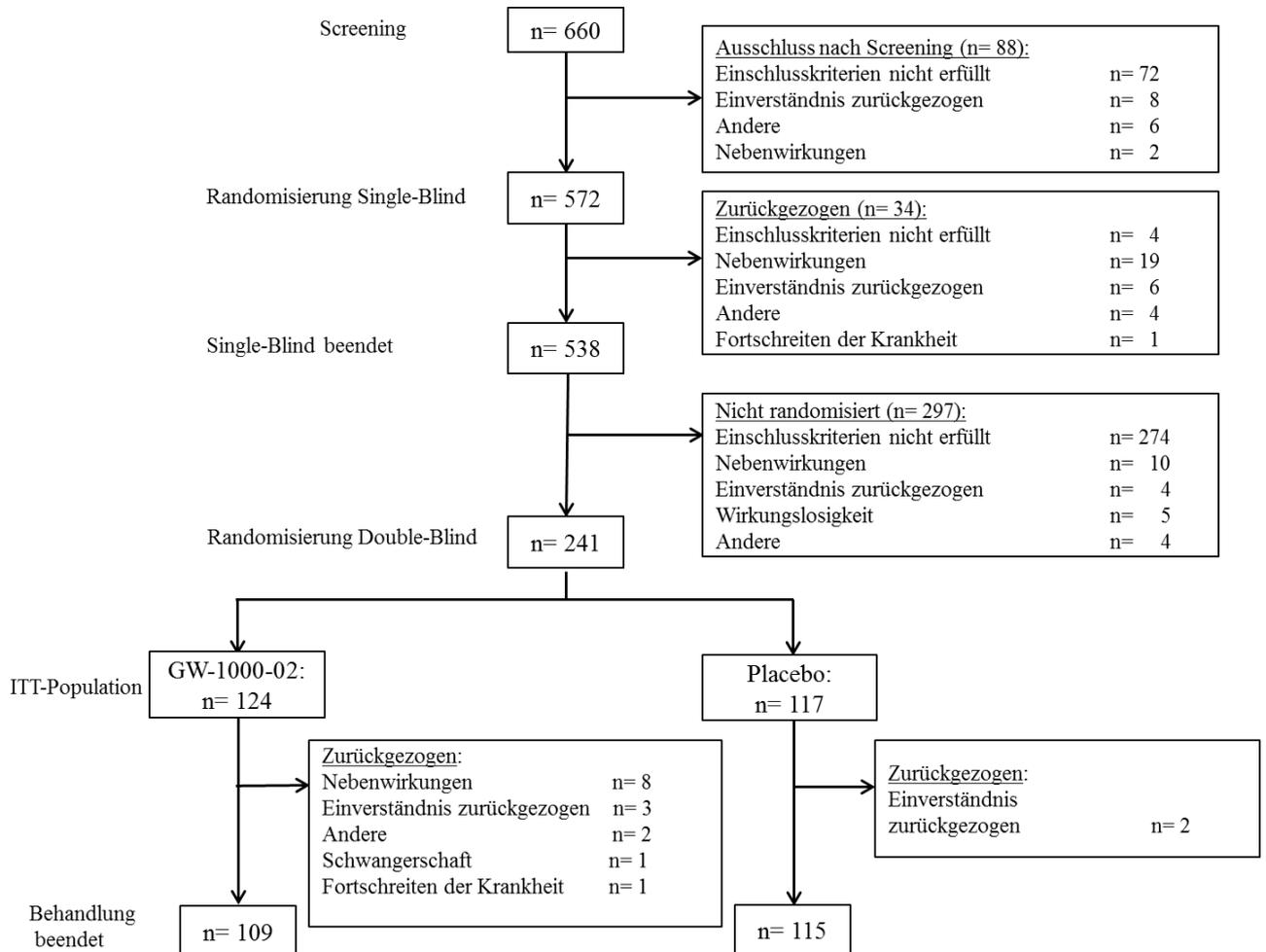


Tabelle 4-95 (Anhang): Studiendesign und -methodik für Studie ARA2009

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Die Studienziele waren: 1. Den Effekt einer 3-wöchigen Behandlung mit einem Cannabinoid Extrakt (THC/CBD etwa 1:1) auf verschiedene psychopathologische Dimensionen zu charakterisieren, einschließlich psychotischer Symptome, Angst und kognitiven Fähigkeiten in Cannabis-naïven Patienten mit MS 2. Die Verträglichkeit und mögliches Suchtpotential zu evaluieren 3. Die Effekte von Cannabis auf die Lebensqualität, Fatigue und Bewegungsfunktionen zu studieren Explorative Studie: Keine statistische Testung einer Hypothese
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Monozentrische, doppelt-blinde, randomisierte, Placebo-kontrollierte, cross-over Studie.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<i>Unbekannt</i>
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	Einschlusskriterien: 1. Männlich oder weiblich, über 18 Jahre alt 3. Diagnostizierte Multiple Sklerose. 4. EDSS Score von 3.5 bis 6.5 5. Bestehende Spastik in mindestens zwei Muskelgruppen 6. Behandlung mit antispastischer Medikation oder Immunmodulatoren 7. Keine Krankheitsgeschichte von Epilepsie, Alkohol- oder Drogenmissbrauch 7. Keine großen Begleiterkrankungen 8. Keine Behandlung mit anti-psychotischer Medikation 9. Keine Konsum von Cannabis vor Studienbeginn
4b	Studienorganisation und Ort der Studiendurchführung	MS Outpatient Clinic "Sapienza" University. Rome Italy
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur	<u>Gruppe:</u> GW-1000-02 (Sativex [®]): Bestandteile: THC (27 mg/ml)/CBD (25 mg/ml) als Extrakt von <i>Cannabis sativa</i> L., mit Pfefferminzöl, 0.05% (v/v), in Ethanol:Propylenglycol (50:50). Applikation in Hüben, jeder Hub setzt 100 µl (THC 2.7 mg und CBD 2.5 mg). <u>Placebo:</u> Bestandteile: Pfefferminzöl, 0.05% (v/v), Chinolingelb, 0.005%

Item ^a	Charakteristikum	Studieninformation
	Administration etc.	(w/v), sunset yellow, 0.0025% (w/v), in Ethanol:Propylenglycol (50:50). Applikation in Hüben, jeder Hub setzt 100 µl Flüssigkeit frei.
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	Beurteilung pathophysiologischer Symptome mittels <i>Symptom Checklist-90</i> und <i>Self-Rating Anxiety Scale</i> Evaluierung der kognitiven Fähigkeiten mittels <i>Paced Auditory Serial Additional Test</i> und <i>Multiple Sclerosis Functional Composite Scale</i> . Lebensqualität und körperliche und psychische Beeinträchtigung mittels <i>VAS</i> , <i>Fatigue Severity Scale</i> und <i>29-item Multiple Sclerosis Impact Scale</i> Beurteilung erfolgte zu Baseline, am Ende der cross-over Periode 1 (Woche 3), am Ende der wash-out Periode (Woche 5) und am Ende der cross-over Periode 2 (Woche 8). Nebenwirkungen
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Unbekannt, aber unwahrscheinlich, da explorative Studie
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	NA – Explorative Studienanlage. Eingeschlossen wurde eine Teilpopulation von geeigneten Patienten aus einer größeren Studie.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	NA
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Computer-generierter Randomisierungs-Code generiert durch einen unabhängigen Statistiker.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Balancierte Randomisierung in den zwei cross-over Armen.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B.	Unbekannt. Beschrieben ist die Ähnlichkeit der beiden Studienmedikationen (Verum und Placebo). Da die Studie monozentrisch durchgeführt wurde, ist von einem geringen Aufwand für eine geeignete Verblindung auszugehen. Die Randomisierung wurde von einem unabhängigen Statistiker durchgeführt.

Item ^a	Charakteristikum	Studieninformation
	nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur 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?	Die Zuteilung erfolgte durch den Investigator konsekutiv bei Patienteneinschluss.
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?	Doppelt-blind. Beide Studienmedikationen (Verum und Placebo) waren weitestmöglich äußerlich ähnlich gestaltet.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Placebo und Verum wurden beide als Spray mit gleicher Färbung verabreicht.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Auswertung erfolgte mit Methoden der deskriptiven Statistik. Ergänzend wurden, je nach Art der untersuchten Variable, Wilcoxon's matched pair signed-rank Test oder t-Test durchgeführt, um potentiell statistisch signifikante Unterschiede auf dem 5 % Niveau zu identifizieren.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	NA
-	Resultate	

Item ^a	Charakteristikum	Studieninformation
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) N = 17 b) Cross-over Gruppe 1: N = 17 / Cross-over Gruppe 2: N = 17 c) N = 17
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	NA - Keine Lost-to-follow-up oder Drop-outs. Alle Patienten, die randomisiert wurden, haben die beiden cross-over Behandlungsphasen durchlaufen und wurden in der statistischen Auswertung berücksichtigt.
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Unbekannt - sehr wahrscheinlich wurde die Studie im Jahr 2004 durchgeführt.
14b	Informationen, warum die Studie endete oder beendet wurde	Hinweise auf eine frühzeitige oder ungeplante Beendigung der Studie liegen nicht vor.
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

NA – Die Studie ARA2009 war eine randomisierte Studie im Cross-over Design: In dieser Studie haben alle Patienten beide Behandlungsgruppen durchlaufen und alle Patienten wurden in die statistische Analyse einbezogen. Es liegt kein Flow-Chart vor.

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.

Tabelle 4-96 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GWMS0001

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:

- 1) **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- 2) **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.

Studie: GWMS0001

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Gwms0001 clinical study report	A
Wade et al. 2004	B

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:

Der Clinical Study Report beschreibt die Studie als:
„[...] randomised [...]“

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:

Folgende Informationen ergeben sich aus dem CSR:

„Randomisation was done using randomised permuted blocks of size four and was stratified by nominated primary impairment and centre.“

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:

Wie die Verdeckung der Gruppenzuteilung erfolgte, ist im CSR wie folgt beschrieben:

„ Study medication was labelled with the Sponsor’s name, study code and subject number [which] were added by pharmacy. The pharmacist at each centre had open access to the randomization schedule, to allocate subjects to study medication. [...] The placebo contained quinolone yellow and sunset yellow to match the colour of the plant extract, and all study medication was provided in amber vials.“

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:

Angaben im CSR:

„[...] double blind [...]“

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:

Angaben im CSR:

„ [...] double blind [...]“

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:

- *Ableich 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:*
 1. *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 2. *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:

Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich, da für jeden definierten Endpunkt vollständige Ergebnisse vorliegen.

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:

Es liegen keine das Verzerrungspotenzial beeinflussende sonstige Aspekte vor.

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:

Das Verzerrungspotential wird als niedrig eingestuft.

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: Veränderung des Grads der Spastizität_**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

Da die Patienten die Endpunkterheber sind, liegt eine Verblindung vor. Dazu steht im CSR: „Subjects recorded in their daily diary the intensity of each impairment [including spasticity] over the past 24 hours on one nominated day each week using a 100 mmm VAS.“

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:

Laut Studienunterlagen (CSR) sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl in der betrachteten Untergruppe (Patienten mit primärer Spastizität) aufgetreten.

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:

Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich, da der Endpunkt zum Studienbeginn definiert und entsprechend ausgewertet wurde.

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

Es gibt keine sonstigen Aspekte, die das Verzerrungspotenzial beeinflussen. Bezüglich des Endpunktes konnten keine relevanten Dateninkonsistenzen zwischen CSR (A) und Publikation (B) festgestellt werden.

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:

Das Verzerrungspotenzial wird als niedrig eingestuft.

Endpunkt: Klinisch relevanter Rückgang der Schwere der Spastizität_

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

Da die Patienten die Endpunkterheber sind, liegt eine Verblindung vor. Dazu steht im CSR:

„Subjects recorded in their daily diary the intensity of each impairment [including spasticity] over the past 24 hours on one nominated day each week using a 100 mm VAS.”

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:

Laut Studienunterlagen (CSR) sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl in der betrachteten Untergruppe (Patienten mit primärer Spastizität) aufgetreten.

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:

Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich, da der Endpunkt zum Studienbeginn definiert und entsprechend ausgewertet wurde.

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:

Es gibt keine sonstigen Aspekte, die das Verzerrungspotenzial beeinflussen.
Bezüglich des Endpunktes konnten keine relevanten Dateninkonsistenzen zwischen
CSR (A) und Publikation (B) festgestellt werden.

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:

Das Verzerrungspotential wird als niedrig eingeschätzt.

Tabelle 4-97 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GWMS0106

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.

Studie: GWMS0106

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
gwms0106-study-report (NCT00711646)	A
Collin et al. 2007	B

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**

X **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:

Punkt 1 und 2 treffen zu

1.

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

X **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:

2:1 Randomisierung Sativex[®]:Placebo; Randomisierung wurde durch einen vorbestimmten Randomisierungscode vorgenommen. Dieser wurde von der GW Statistics Abteilung produziert mit einer Behandlungszuteilung in Dreier-Blöcken (A)(B).

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“)**

X **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:

Studienmedikation wurde von GW Clinical Trials Supplies Department eingepackt und mit den Teilnehmernummern versehen (A). Die Studienmedikation wurde in 5,5ml Typ I bernsteinfarbenen Glasfläschchen verschickt, mit dem GW-Namen, Studiencode, Teilnehmernummer, Visitennummern, Packungsreferenznummer und dem Verfallsdatum darauf (A)

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

X **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:

Doppelt-blindes Studiendesign (A)(B)

Behandler bzw. Weiterbehandler

X **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:

Doppelt-blindes Studiendesign (A, B)

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:

- *Ableich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Ableich 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.

X **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:

Darstellung der Ergebnisse für ITT Gesamtpopulation. Die Ergebnisdarstellung erfolgte unabhängig von der statistischen Signifikanz der Ergebnisse. Analyse und Darstellung der Ergebnisse wurden wie im Protokoll beschrieben durchgeführt.

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:

Der Patientenfluss wurde transparent dargestellt. Begleitbehandlungen in den Gruppen sind dargestellt und vergleichbar.

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.

X 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: Schwere der Spastik auf NRS (Absolute Veränderung und Anzahl Patienten mit klinisch relevanter Verbesserung)

5. 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.

Xja: 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

Doppelt-blindes Design der Studie. Der Schweregrad der Spastik wurde durch die Patienten erhoben, die verblindet waren.

6. 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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

7. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

8. **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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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 der Spastiken pro Tag

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.

Xja: 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

Doppelt-blindes Design der Studie. Die Spastik wurde durch die Patienten, die verblindet waren, im Tagebuch dokumentiert.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden wie im Protokoll geplant durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Klinische Globalbeurteilung der Veränderung durch den Pateinten (SGIC)

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.

Xja: 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

Doppelt-blindes Studiendesign. Der Endpunkt wurde von den Patienten erhoben, die verblindet waren.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Nebenwirkungen**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.

Xja: 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

Doppelt-blindes Studiendesign. Der Endpunkt wurde von den Investigatoren dokumentiert, die verblindet waren.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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:

Tabelle 4-98 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GWCL0403

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.

Studie: GWCL0403

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
GWCL0403-study-report [EudraCT Nummer: 2004-002509-63]	A
Collin et al. 2010	B

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:

Punkt 1 und 2 treffen zu.

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:

Computerbasierte Randomisierung (A)

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:

Studienmedikation wurde von GW Clinical Trials Supplies Department eingepackt und mit den Teilnehmernummern versehen (A). Die Studienmedikation wurde in 5,5 ml Typ I bernsteinfarbenen Glasfläschchen verschickt, mit dem GW-Namen, Studiencode, Teilnehmernummer, Visitennummern, Packungsreferenznummer und dem Verfallsdatum darauf (A).

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

X **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:

Doppelt-blind Studie (A, B)

Behandler bzw. Weiterbehandler

X **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:

Doppelt-blind Studie (A, B)

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:

- 3. erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - 4. 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.

X 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:

Darstellung der Ergebnisse für ITT Gesamtpopulation. Die Ergebnisdarstellung erfolgte unabhängig von der statistischen Signifikanz der Ergebnisse. Analyse und Darstellung der Ergebnisse wurden, wie im Protokoll beschrieben, durchgeführt.

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:

Der Patientenfluss wurde transparent dargestellt. Begleitbehandlungen in den Gruppen sind dargestellt und vergleichbar.

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.

X 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:

Endpunkt: Schwere der Spastik auf NRS (Absolute Veränderung und Anzahl Patienten mit klinisch relevanter Verbesserung)

(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.

Xja: 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

Doppelt-blindes Design der Studie. Der Schweregrad der Spastik wurde durch die Patienten erhoben, die verblindet waren.

(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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

(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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

(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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen
wurden nicht festgestellt

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.

X 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: Gesundheitszustand (EQ-5D)**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.

Xja: 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

Doppelt-blindes Studiendesign. Der Endpunkt wurde von den Patienten dokumentiert.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Lebensqualität**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.

Xja: 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

Doppelt-blindes Studiendesign. Der Endpunkt wurde von den Patienten dokumentiert.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Schlafqualität

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.

Xja: 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

Doppelt-blindes Studiendesign. Der Endpunkt wurde von den Patienten dokumentiert, die verblindet waren.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Aktivitäten des täglichen Lebens (Barthel Index)**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.

Xja: 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

Doppelt-blindes Studiendesign. Der Endpunkt wurde von den Investigatoren dokumentiert, die verblindet waren.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Gehfähigkeit**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.

Xja: 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

Doppelt-blindes Studiendesign. Der Endpunkt wurde von den Ärzten erhoben, die verblindet waren.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Klinische Globalbeurteilung der Veränderung durch den Betreuenden (CGIC)

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.

Xja: 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

Doppelt-blindes Studiendesign. Der Endpunkt wurde von den betreuenden Patienten erhoben, die verblindet waren.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Nebenwirkungen**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.

Xja: 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

Doppelt-blindes Studiendesign. Der Endpunkt wurde von den Investigatoren dokumentiert, die verblindet waren.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

3.) **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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:

Tabelle 4-99 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie GWSP0604

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.

Studie: GWSP0604

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
GWSP0604-study-report [EudraCT Number: 2006-005910-11]	A
Novotna et al. 2011	B

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:

Punkt 1 und 2 treffen zu.

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:

Computerbasierte Randomisierung (A)

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:

Zuteilung über Interactive Voice Response System (A)

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- 1) Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- 2) 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

X **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:

Phase A: Einfache Verblindung (Patienten), alle Patienten erhielten Verum,
Phase B: Doppelt-blind randomisierte Gruppenzuteilung (A, B)

Behandler bzw. Weiterbehandler

X **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:

Phase B: Doppelt-blind randomisierte Gruppenzuteilung (A, B)

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:

 5. *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 6. *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.

X 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:

Darstellung der Ergebnisse für ITT Gesamtpopulation. Die Ergebnisdarstellung erfolgte unabhängig von der statistischen Signifikanz der Ergebnisse. Analyse und Darstellung der Ergebnisse wurden, wie im Protokoll beschrieben, durchgeführt.

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:

Der Patientenfluss wurde transparent dargestellt. Begleitbehandlungen in den Gruppen sind dargestellt und vergleichbar.

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.

X 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:

Endpunkt: Schwere der Spastik auf NRS (Absolute Veränderung und Anzahl Patienten mit klinisch relevanter Verbesserung)

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.

Xja: 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 Schweregrad der Spastik wurde durch die Patienten erhoben, die in allen Studienphasen (einarmige, einfach verblindete Run-in Phase A, parallele, doppelt-blinde, randomisierte Phase B) verblindet waren.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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 der Spastiken pro Tag

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.

Xja: 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 Endpunkt wurde durch die Patienten erhoben, die in allen Studienphasen (einarmige, einfach verblindete Run-in Phase A, parallele, doppelt-blinde, randomisierte Phase B) verblindet waren.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Gesundheitszustand (EQ-5D)**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.

Xja: 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 Endpunkt wurde durch den Patienten erhoben, die in allen Studienphasen (einarmige, einfach verblindete Run-in Phase A, parallele, doppelt-blinde, randomisierte Phase B) verblindet waren.

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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Lebensqualität**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.

Xja: 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

Doppelt-blindes Studiendesign. Der Endpunkt wurde von den Patienten dokumentiert.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Schlafqualität

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.

Xja: 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 Endpunkt wurde bei Patienten erhoben, die in allen Studienphasen (einarmige, einfach verblindete Run-in Phase A, parallele, doppelt-blinde, randomisierte Phase B) verblindet waren.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Aktivitäten des täglichen Lebens (Barthel Index)**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.

Xja: 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

In der einfach-verblindeten Run-in Phase erhielten alle Patienten Verum. Die anschließende parallele randomisierte Phase wurde doppelt-blind durchgeführt, d.h. auch der Investigator als Endpunkterheber war verblindet.

• 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.

Xja: Eines der folgenden Merkmale trifft zu:

- 3) 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.
- 4) 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).
- 5) 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

- **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

- **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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Gehfähigkeit (10m-Gehstreckentest)**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.

Xja: 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

In der einfach-verblindeten Run-in Phase erhielten alle Patienten Verum. Die anschließende parallele randomisierte Phase wurde doppelt-blind durchgeführt, d.h. auch der Investigator als Endpunkterheber war verblindet.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Klinische Globalbeurteilung der Veränderung durch den Patienten (SGIC)

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.

Xja: 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 Endpunkt wurde bei Patienten erhoben, die in allen Studienphasen (einarmige, einfach verblindete Run-in Phase A, parallele, doppelt-blinde, randomisierte Phase B) verblindet waren.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden wie im Protokoll geplant durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Klinische Globalbeurteilung der Veränderung durch die betreuende Person (CGIC)

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.

Xja: 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 Endpunkt wurde von den betreuenden Personen dokumentiert, die keine Kenntnisse zur Studienbehandlung hatten.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Klinische Globalbeurteilung der Veränderung durch den Arzt (PhGIC)

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.

Xja: 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

In der einfach-verblindeten Run-in Phase erhielten alle Patienten Verum. Die anschließende parallele randomisierte Phase wurde doppelt-blind durchgeführt, d.h. auch der Investigator als Endpunkterheber war verblindet.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Nebenwirkungen**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.

Xja: 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

In der einfachverblindeten Run-in Phase erhielten alle Patienten Verum. Die anschließende parallele randomisierte Phase wurde doppelt-blind durchgeführt, d.h. auch der Investigator als Endpunkterheber war verblindet.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

X 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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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:

Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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:

Tabelle 4-100 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie ARA2009

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.

Studie: ARA2009

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Publikation Aragona et al 2009	A

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:

Punkt 1 und 2 treffen bedingt zu.

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:

Computer-generated Code generiert durch einen unabhängigen Statistiker. Die Randomisierung erfolgte balanciert in die beiden Behandlungsgruppen. (A)

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

Xunklar: 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:

Fehlende Beschreibung zur Verdeckung der Gruppenzuteilung. Jedoch Beschreibung von Verum und Placebo Medikation, so dass hier von äußerlich identischer Studienmedikation ausgegangen werden kann (A). Zudem war die Studie monozentrisch angelegt und durchgeführt, was eine einfache Logistik zur Verblindung und Zuteilung ermöglicht haben sollte. (A)

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

X **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:

Doppelt-blind Studie (A)

Behandler bzw. Weiterbehandler

X **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:

Doppelt-blind Studie (A)

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.

X 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:

Darstellung der Ergebnisse für ITT Gesamtpopulation. Eine ergebnisabhängige Darstellung ist nicht erkennbar. Die Ergebnisdarstellung erfolgte unabhängig von der statistischen Signifikanz der Ergebnisse. (A)

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:

Durch das cross-over Design und die relativ kurzen Behandlungsphasen (je 3 Wochenbehandlungszeit mit einmal 2 Wochen Wash-out) und der Gesamtstudiendauer von 8 Wochen sind die Behandlungsgruppen vergleichbar, zumal für alle 17 randomisierten Patienten Daten aus beiden Behandlungsphasen vorliegen. (A)

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.

X 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:

Endpunkt: Lebensqualität

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.

Xja: 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

Doppelt-blindes Studiendesign. Der Endpunkt wurde von den Patienten dokumentiert.

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs wurden transparent mit Gründen beschrieben.

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:

Die Analysen wurden, wie im Protokoll geplant, durchgeführt, zusätzlich durchgeführte Analysen wurden begründet und alle Ergebnisse sind im Clinical Study Report dargestellt.

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

 Relevante Dateninkonsistenzen oder unplausible Angaben zwischen Quellen wurden nicht festgestellt.

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.

X 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: Nebenwirkungen**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.

Xja: 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

Doppelt-blindes Studiendesign. Der Endpunkt wurde von den Investigatoren dokumentiert, die verblindet waren. (A)

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.

Xja: 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:

Auswertung der ITT Population, Lost-to-follow-up und Drop-outs gab es keine (A). Alle 17 randomisierten Patienten haben beide Behandlungsphasen durchlaufen. (A)

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:

Hierzu liegen keine Hinweise vor. Alle aufgetretenen Unerwünschten Ereignisse wurden analysiert und dargestellt. (A)

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:

Relevante Dateninkonsistenzen oder unplausible Angaben in den Ergebnissen wurden nicht festgestellt. (A)

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.

X 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:
