



**Alemtuzumab, Cladribin,
Dimethylfumarat, Fingolimod,
Natalizumab, Ocrelizumab,
Ofatumumab, Ozanimod,
Ponesimod und Teriflunomid
zur Behandlung Erwachsener
mit hochaktiver schubförmig
remittierender multipler
Sklerose**

Berichtsplan

Auftrag: A20-60
Version: 2.1
Stand: 03.12.2021

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Alemtuzumab, Cladribin, Dimethylfumarat, Fingolimod, Natalizumab, Ocrelizumab, Ofatumumab, Ozanimod, Ponesimod und Teriflunomid zur Behandlung Erwachsener mit hochaktiver schubförmig remittierender multipler Sklerose

Auftraggeber

Gemeinsamer Bundesausschuss

Datum des Auftrags

16.07.2020

Interne Auftragsnummer

A20-60

Anschrift des Herausgebers

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8
50670 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

Schlagwörter

Alemtuzumab, Cladribin, Dimethylfumarat, Fingolimod, Natalizumab, Ocrelizumab, Ofatumumab, Ozanimod, Ponesimod, Teriflunomid, Multiple Sklerose – Schubförmige Remittierende, Nutzenbewertung, Systematische Übersicht

Keywords

Alemtuzumab, Cladribine, Dimethyl Fumarate, Fingolimod Hydrochloride, Natalizumab, Ocrelizumab, Ofatumumab, Ozanimod, Ponesimod, Teriflunomide, Multiple Sclerosis – Relapsing-Remitting, Benefit Assessment, Systematic Review

Inhaltsverzeichnis

| | Seite |
|--|-----------|
| Tabellenverzeichnis | iv |
| Abkürzungsverzeichnis | v |
| 1 Hintergrund | 1 |
| 2 Fragestellung | 8 |
| 3 Projektverlauf | 9 |
| 3.1 Zeitlicher Verlauf des Projekts | 9 |
| 3.2 Dokumentation der Änderungen im Projektverlauf | 10 |
| 4 Methoden | 12 |
| 4.1 Kriterien für den Einschluss von Studien in die Untersuchung | 12 |
| 4.1.1 Population..... | 12 |
| 4.1.2 Prüf- und Vergleichsintervention | 12 |
| 4.1.3 Patientenrelevante Endpunkte | 13 |
| 4.1.4 Studientypen | 14 |
| 4.1.5 Studiendauer | 14 |
| 4.1.6 Publikationssprache | 14 |
| 4.1.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss | 14 |
| 4.1.8 Vorgehen im Falle einer Zulassungsänderung im Projektverlauf | 15 |
| 4.1.9 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen..... | 15 |
| 4.2 Informationsbeschaffung | 15 |
| 4.2.1 Informationsquellen..... | 16 |
| 4.2.2 Selektion relevanter Studien..... | 17 |
| 4.3 Informationsbewertung und Synthese | 17 |
| 4.3.1 Darstellung der Einzelstudien..... | 17 |
| 4.3.2 Bewertung des Verzerrungspotenzials der Ergebnisse..... | 18 |
| 4.3.3 Netzwerk-Metaanalysen | 19 |
| 4.3.4 Sensitivitätsanalysen | 22 |
| 4.3.5 Subgruppenmerkmale und andere Effektmodifikatoren | 22 |
| 4.3.6 Aussagen zur Beleglage | 23 |
| 5 Literatur | 24 |

Tabellenverzeichnis

| | Seite |
|--|--------------|
| Tabelle 1: Fragestellungen der Nutzenbewertung..... | 8 |
| Tabelle 2: Übersicht über die Kriterien für den Studieneinschluss..... | 15 |

Abkürzungsverzeichnis

| Abkürzung | Bedeutung |
|------------------|---|
| AWMF | Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V. |
| CONSORT | Consolidated Standards of Reporting Trials |
| G-BA | Gemeinsamer Bundesausschuss |
| Gd+ | gadoliniumanreichernd |
| EDSS | Expanded Disability Status Scale |
| ICH | International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use |
| IQWiG | Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen |
| ITT | Intention to treat |
| KI | Konfidenzintervall |
| KKNMS | Krankheitsbezogenes Kompetenznetz multiple Sklerose e. V. |
| MAGNIMS | Magnetic Resonance Imaging in Multiple Sclerosis |
| MRT | Magnetresonanztomografie |
| MSFC | Multiple Sclerosis Functional Composite |
| MS | multiple Sklerose |
| NMA | Netzwerk-Metaanalyse |
| PML | progressive multifokale Leukenzephalopathie |
| PPMS | primär progrediente multiple Sklerose |
| PRO | Patient-reported Outcome (patientenberichteter Endpunkt) |
| RCT | Randomized controlled Trial (randomisierte kontrollierte Studie) |
| RMS | Relapsing multiple Sclerosis (schubförmige multiple Sklerose) |
| RRMS | Relapsing-remitting multiple Sclerosis (schubförmig remittierende multiple Sklerose) |
| SMD | Standardized Mean Difference (standardisierte Mittelwertdifferenz) |
| SPMS | sekundär progrediente multiple Sklerose |
| SUE | schwerwiegendes unerwünschtes Ereignis |
| UE | unerwünschtes Ereignis |
| ZNS | Zentralnervensystem |

1 Hintergrund

Epidemiologie und Verlauf der multiplen Sklerose

Die multiple Sklerose (MS) ist mit einer Prävalenz von über 200 000 Erkrankten in Deutschland eine der häufigsten neuroinflammatorisch-neurodegenerativen Erkrankungen [1]. Für die Betroffenen bedeutet eine MS-Erkrankung oft lebenslange und weitreichende Konsequenzen für den Alltag, da sie neben der Akutsymptomatik (z. B.: motorische Störungen, Sehstörungen, Schmerzen, Inkontinenz) in vielen Fällen zu einer fortschreitenden Behinderung bis hin zur Pflegebedürftigkeit führt [2-4]. Zudem treten häufig „unsichtbare“, aber für die Betroffenen belastende Symptome wie Fatigue, Depressionen oder kognitive Einschränkungen auf, die die Aktivitäten des täglichen Lebens erheblich einschränken. Die Erkrankung tritt in der Regel zwischen dem 20. und 40. Lebensjahr auf. Frauen sind 2- bis 3-mal häufiger von MS betroffen als Männer [5].

Die MS tritt im Wesentlichen in 3 Verlaufsformen auf. Neben der primär progredienten MS (PPMS), die insgesamt selten (< 10 %) und durch eine Behinderungsprogression von Erkrankungsbeginn an gekennzeichnet ist, tritt bei der großen Mehrheit der Betroffenen die schubförmig remittierende Verlaufsform (Relapsing-remitting multiple Sclerosis, RRMS) auf [2]. Diese ist gekennzeichnet durch wiederkehrende Schübe. Darunter wird das Auftreten von Episoden mit MS-typischer klinischer Symptomatik verstanden, die über einen Zeitraum von wenigstens 24 Stunden auftreten und auf fokale oder multifokale entzündlich-demyelinisierende Ereignisse im Zentralnervensystem (ZNS) zurückzuführen sind und nicht durch eine Erhöhung der Körpertemperatur (Uhthoff-Phänomen) oder Infektionen verursacht werden [6]. Intensität, Zeitpunkt des Auftretens, Dauer und Art der Symptome sind dabei individuell verschieden ausgeprägt. Mit dem Ende eines Schubes klingen bei den Betroffenen häufig die durch den Schub verursachten Einschränkungen wieder ab. Teilweise hinterlassen die Schübe aber auch Residuen, die zu einer bleibenden Behinderung oder Verschlechterung einer bestehenden Behinderung führen. Zudem kommt es auch bei RRMS mitunter zu einer schubunabhängigen Behinderungsprogression [7]. Bei mindestens 50 % der an RRMS Erkrankten geht die Erkrankung irgendwann in eine sekundär progrediente MS-Verlaufsform (SPMS) über. Dann tritt eine stetige Behinderungsprogression analog zur PPMS ein, während der weitere Schübe auftreten können [8]. Darüber hinaus werden sowohl bei der RRMS als auch bei den progredienten Formen aktive und nicht aktive Verläufe unterschieden, wobei die Aktivität durch das Auftreten von Schüben oder Läsionen in der Magnetresonanztomografie (MRT) gekennzeichnet ist [9].

Das Symptomspektrum ist breit und kann u. a. Bewegungsstörungen (Ataxie, Tremor, Spastik), Blasen- und Darmstörungen, Sehstörungen, sexuelle Störungen und Schmerzen umfassen, insbesondere aber auch neuropsychologische Symptome wie Fatigue, Schlafstörungen, kognitive Einschränkungen und Depressionen [2]. Neben neurologischer Behinderung ist insbesondere Fatigue eine signifikante Ursache für eine Einschränkung der Aktivitäten des täglichen Lebens sowie für Arbeitsunfähigkeit und Frühverrentung [10]. Depressionen beispielsweise treten unter MS-Erkrankten häufiger auf als in der Gesamtbevölkerung und sind

pathophysiologisch vermutlich auf die MS zurückzuführen [11,12]. Es sind vor allem diese neuropsychiatrischen und neuropsychologischen Symptome, die das psychosoziale Wohlbefinden sowie die Funktionen und Aktivitäten des Alltags und des Berufslebens der Betroffenen einschränken. Neben der häufig auftretenden Fatigue werden von Betroffenen Sehstörungen, kognitive Einschränkungen und motorische Störungen häufig als besonders relevante Symptome genannt [13].

Da sich die vorliegende Nutzenbewertung gemäß Auftrag ausschließlich mit der Verlaufsform der hochaktiven RRMS befasst, wird auf die progredienten Verlaufsformen und die nicht hochaktive RRMS im Folgenden nicht weiter eingegangen.

Definition der hochaktiven RRMS

Eine allgemein anerkannte Definition hochaktiver RRMS existiert derzeit nicht. Darüber hinaus existieren weitere verschiedene Begrifflichkeiten, z. B. aggressive und maligne MS, die ebenfalls unscharf definiert sind und jeweils eine Teilmenge der RRMS abbilden [14-16]. Auch in der aktuellen S2-Leitlinie der Deutschen Gesellschaft für Neurologie erfolgt eine Definition der hochaktiven RRMS lediglich für therapienaive Patientinnen und Patienten [5].

In den Zulassungsstudien der hier untersuchten Wirkstoffe beruht die Definition einer hochaktiven RRMS im Wesentlichen auf der Schubrate und / oder der Anzahl neuer Läsionen in der MRT. Die Patientenpopulationen dieser Studien sind häufig nicht auf die hochaktive Ausprägung beschränkt (z. B. Natalizumab, Cladribin, Alemtuzumab). In einigen Fällen wurden jedoch seitens der Zulassungsbehörden Einschränkungen vorgenommen, entweder um Sicherheitsbedenken Rechnung zu tragen und / oder aufgrund geringerer Wirksamkeit bei Patienten ohne aktive Erkrankung. Diese Präparate wurden darum in der Anwendung auf Patientinnen und Patienten mit aktiver oder hochaktiver Erkrankung beschränkt. Die Operationalisierung der Aktivität in klinischen Studien variiert jedoch [17]. In der Regel wird von mindestens 1 oder 2 Schüben im vorangegangenen Jahr ausgegangen zuzüglich mehrerer neuer T2-Läsionen und / oder mindestens 1 gadoliniumanreichernden (Gd+) T1-Läsion trotz angemessener krankheitsmodifizierender Therapie. Daneben existieren auch Operationalisierungen hoher Krankheitsaktivität, die auf der Behinderungsprogression oder dem Grad der Behinderung bei der Diagnose beruhen, i. d. R. anhand der Extended Disability Status Scale (EDSS) bestimmt [14,16]. Eine explizite Definition des Anwendungsgebiets, sofern eine Beschränkung auf hochaktive RRMS vorliegt, erfolgt in den Fachinformationen der einzelnen Wirkstoffe nicht [18-28].

Für die vorliegende Bewertung wird eine Operationalisierung der Krankheitsaktivität bevorzugt, die auf patientenrelevanten Merkmalen beruht [29]. Das Auftreten von Schüben ist ein solches Merkmal. In der Literatur finden sich Beispiele für eine Charakterisierung hoher Krankheitsaktivität ausschließlich anhand der Schubraten [30-32]. Spelman 2020 zeigte, dass eine Operationalisierung hoher Aktivität, die auf Schubraten beruht, eine höhere Assoziation mit Behinderungsprogression zeigt als eine MRT-basierte Operationalisierung [31].

Mehrere Studien zeigten, dass unter Behandlung mit Interferonen oder Glatirameracetat die Ausbildung mehrerer neuer T2-Läsionen (≥ 3) prognostisch für einen ungünstigen Krankheitsverlauf ist [33-35]. Daher wird zur Beurteilung der Krankheitsaktivität zusätzlich zum Auftreten von Schüben auch die MRT-Aktivität betrachtet. Läsionen sind 5- bis 10-mal häufiger als Schübe, für RRMS-Betroffene jedoch nicht zwingend als klinische Symptomatik spürbar. Zudem ist ihre Erfassung mit relevanten Unsicherheiten behaftet, denn die Aussagekraft der MRT-Diagnostik hängt u. a. vom verwendeten Protokoll und dem Umgang mit Kontrastmitteln ab, und auch davon, ob alle betroffenen ZNS-Regionen erfasst werden [33]. Gd⁺-Läsionen sind nur über einen kurzen Zeitraum darstellbar und daher nicht sensitiv genug, um das inflammatorische Geschehen über einen längeren Zeitraum zu beurteilen. MRTs unter Einsatz von Gadolinium als Kontrastmittel sind für Patientinnen und Patienten möglicherweise mit einem höheren Risiko behaftet im Vergleich zu Untersuchungen ohne zusätzliche Kontrastmittelbeigabe [36,37]. Inwieweit eine kurzfristige Verschlechterung des EDSS-Scores unter Therapie prädiktiv für eine langfristige, irreversible Behinderungsprogression ist, scheint unklar zu sein, zumindest wird sie für Scoringssysteme wie den modifizierten Rio- oder den Magnetic-Resonance-Imaging-in-MS(MAGNIMS)-Score diskutiert, aber nicht durchgängig herangezogen [38,39].

Trotz der beschriebenen Unsicherheiten ist bei Vorliegen einer sehr hohen MRT-Aktivität davon auszugehen, dass zumindest verdeckt klinische Symptome aufgetreten sind, jedoch möglicherweise nicht dokumentiert oder diagnostiziert wurden. Für Studien, in die Patientinnen und Patienten eingeschlossen wurden, bei denen Schubereignisse unter Vortherapie nicht erfasst wurden, muss die Vergleichbarkeit der untersuchten Populationen gewährleistet sein. Dazu muss die MRT-Erfassung zwischen den Studien vergleichbar, d. h. nach anerkannten Standards und Protokollen erfolgt sein [33,40] und adäquat berichtet werden. Innerhalb einer Studie müssen alle MRT-Scans longitudinal mit denselben Sequenzen und Sequenzparametern durchgeführt worden sein, damit auch eine intraindividuelle Vergleichbarkeit der Scans bei einzelnen Patientinnen und Patienten gewährleistet ist. Unter diesen Voraussetzungen ist auch eine Operationalisierung der hochaktiven Erkrankung denkbar, die ausschließlich auf einer hohen MRT-Aktivität basiert.

Es existiert derzeit keine allgemein anerkannte Definition allein anhand der MRT-Aktivität, welche für die Operationalisierung einer hochaktiven Erkrankung geeignet ist. Im Rahmen des Zulassungsverfahrens zu Natalizumab wurde ein Schwellenwert von 9 T2-Läsionen näher untersucht [41]. Darüber hinaus wird das Auftreten von ≥ 9 neuen oder vergrößerten T2-Läsionen in den Fachinformationen weiterer Wirkstoffe in der Definition der hochaktiven Erkrankung berücksichtigt [19,20]. Diese Schwelle kann zum Zwecke der Nutzenbewertung daher als Richtwert für eine rein MRT-basierte Operationalisierung von hochaktiver RRMS dienen. Für die Nutzenbewertung wird dabei nicht zwischen T1- und T2-Läsionen unterschieden (im Sinne von Combined Unique Active Lesions). In den aktuellen MAGNIMS-Empfehlungen [42] wird die Aufnahme von T1-Läsionen zur Bestimmung der Krankheitsaktivität als optional eingestuft. Der Gebrauch von Gadolinium als Kontrastmittel

wird zudem nur unter bestimmten Bedingungen empfohlen, z. B. wenn die Detektion von T2-Läsionen allein nicht ausreicht.

Zusammenfassend werden für die vorliegende Bewertung verschiedene Definitionen einer hochaktiven RRMS berücksichtigt, die auf dem Auftreten von Schüben und / oder der MRT-Aktivität basieren. Dabei sind ausschließlich klinische, ausschließlich MRT-basierte oder daraus kombinierte Definitionen anwendbar:

- Anhand ausschließlich klinischer Parameter kann von einer hochaktiven Erkrankung ausgegangen werden, wenn ≥ 1 Schub in den letzten 12 Monaten oder ≥ 2 Schübe in den letzten 24 Monaten aufgetreten sind, die jeweils mit einer erheblichen funktionellen Beeinträchtigung (z. B. Visusverschlechterung, Verschlechterung einzelner EDSS-Komponenten) einhergingen. Hierbei ist aufgrund der klinischen Relevanz des Schubgeschehens eine zusätzliche bildgebende Anamnese nicht erforderlich.
- Wenn Schübe auftreten, die nicht zu erheblichen funktionellen Beeinträchtigungen führen, kann unterstützend das Auftreten neuer oder vergrößerter Läsionen als Kriterium hinzugezogen werden. Das Auftreten von ≥ 1 Schub in den letzten 12 Monaten und zusätzlich ≥ 3 neuen oder vergrößerten T2-Läsionen oder ≥ 1 neuen Gd⁺-Läsion in einem Verlaufs-MRT kann hier als Annäherung dienen.
- Eine rein MRT-basierte Definition ohne Berücksichtigung klinischer Manifestationen kann auf einem Richtwert von ≥ 9 neuen oder vergrößerten Läsionen innerhalb der letzten 12 Monate beruhen.

Grundsätzlich gilt bei allen Definitionen, dass die Bewertung der Krankheitsaktivität über das Auftreten von Schüben oder MRT-Aktivität erst bei voller Wirksamkeit einer ausreichenden und angemessenen Vortherapie erfolgen sollte. Als vollständige und angemessene Vorbehandlung wird eine krankheitsmodifizierende Therapie unter voller Wirksamkeit des jeweiligen Wirkstoffes angesehen. Dabei sind die Latenzzeiten der verabreichten Wirkstoffe zu beachten. Als Vortherapie kommen insbesondere Interferon beta, Glatirameracetat, Teriflunomid und Dimethylfumarat in Betracht. Interferon beta, Teriflunomid und Dimethylfumarat erreichen ihre volle Wirksamkeit nach ca. 3 Monaten kontinuierlicher Therapie, Glatirameracetat nach ca. 6 Monaten [33,43]. Unter der Voraussetzung, dass die Wirkstoffe in ausreichender und stabiler Dosierung verabreicht wurden, ergeben sich damit ca. 3 bis 6 Monate Therapiedauer für eine vollständige und angemessene Vortherapie. Schubereignisse oder neu auftretende Läsionen, die in diese Latenzzeit fallen, haben für die Bewertung der Wirksamkeit nur begrenzte Aussagekraft, und die Beurteilung der Schubaktivität sowie der in der MRT nachgewiesenen Veränderungen ist dementsprechend auf den Zeitraum nach Erreichen der vollen Wirksamkeit zu beziehen. Für die Beurteilung der MRT-Aktivität ist daher eine Neuerfassung des Läsionsstatus nach Erreichen der vollen Wirksamkeit erforderlich (Re-Baselining) [42]. Die Feststellung einer hochaktiven RRMS erfolgt dann nach Vollenden einer vollständigen immunmodulierenden Therapie.

Die oben beschriebenen Definitionen für die hochaktive RRMS sind grundsätzlich von einer lediglich aktiven Erkrankung bei Nichtansprechen auf eine angemessene und ausreichende Vortherapie zu unterscheiden. Eine aktive Erkrankung kann bereits bei Auftreten 1 neuen T2- oder T1-Läsion oder einem Schub in einem bestimmten Zeitraum gegeben sein [44].

Abzugrenzen ist die hochaktive RRMS von der SPMS, die mit aufgesetzten Schüben einhergehen kann. Eine exakte Abgrenzung ist schwierig, da die Übergänge fließend sind und keine exakten Biomarker existieren. Die Diagnose der SPMS erfolgt hauptsächlich nach Lebensalter der Betroffenen und der Krankheitsdauer. Sie ist daher auch nur retrospektiv anhand des klinischen Verlaufs sicher zu diagnostizieren [9]. SPMS-Formen mit aufgesetzten Schüben sind dabei eher dem Spektrum der schubförmigen MS zuzuordnen, da die Zunahme der Behinderung an unvollständig remittierte Schübe gekoppelt sein kann [45].

Ziele der Behandlung

Die Behandlung der MS umfasst im Wesentlichen die Therapie akuter Schübe, die Behandlung der Krankheitssymptomatik und die Modifikation des Krankheitsverlaufs, also Behinderungsprogression und Schubrate. Zur Schubtherapie werden in der Regel Glukokortikoide, vor allem Methylprednisolon, eingesetzt, um Symptome zu mildern und die Schubdauer zu verkürzen. Parallel erfolgt schubunabhängig eine dauerhafte krankheits- bzw. verlaufsmodifizierende Therapie mit dem Ziel, die Schubfrequenz zu verringern und eine Behinderungsprogression zu verhindern oder zu verzögern. Hierzu werden immunmodulatorische Wirkstoffe eingesetzt [2,46-48].

Seit den 1990er-Jahren werden Interferon beta-1a und -1b sowie Glatirameracetat als immunmodulatorische Therapien eingesetzt [49]. Mit den Zulassungen von Alemtuzumab, Cladribin, Dimethylfumarat, Fingolimod, Natalizumab, Ocrelizumab, Ofatumumab, Ozanimod, Ponesimod und Teriflunomid stehen seit 2005 weitere immunmodulatorische Wirkstoffe zur Verfügung. Diese sind Gegenstand der vorliegenden Bewertung.

Die Therapie der MS erfolgte ursprünglich nach einem Stufentherapieschema, welches zwischen Basistherapien und Eskalationstherapien bei ungenügendem Ansprechen bzw. bei schweren Verläufen unterschied [46]. Zu den Basistherapeutika zählten dabei insbesondere Interferone und Glatirameracetat, aber auch Teriflunomid und Dimethylfumarat. Die aktuelle S2-Leitlinie der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) teilt die verfügbaren Wirkstoffe inzwischen in Wirksamkeitskategorien ein. Für nicht hochaktive Verläufe werden dabei insbesondere Interferone, Glatirameracetat, Dimethylfumarat und Teriflunomid empfohlen, die jetzt der Wirksamkeitskategorie 1 entsprechen [5]. Das Krankheitsbezogene Kompetenznetz Multiple Sklerose e. V. (KKNMS) empfiehlt neben Interferonen und Glatirameracetat ebenfalls den Einsatz von Dimethylfumarat und Teriflunomid bei milden und moderaten Verläufen, allerdings sind beide grundsätzlich auch bei hochaktiver MS bzw. als Eskalationstherapie einsetzbar [19,20,46]. Die hier betrachteten neuen Wirkstoffe stellen damit überwiegend Eskalationstherapien dar, die bei Versagen einer Vorbehandlung und anhaltend (hoch) aktivem Krankheitsverlauf eingesetzt

werden. Ihre Zulassungen sind nicht deckungsgleich und umfassen das hier betrachtete Anwendungsgebiet ganz oder teilweise.

Gegenstand der vorliegenden Bewertung ist der Vergleich der Eskalationstherapien Alemtuzumab, Cladribin, Dimethylfumarat, Fingolimod, Natalizumab, Ocrelizumab, Ofatumumab, Ozanimod, Ponesimod und Teriflunomid. Andere Wirkstoffe wie Interferone oder Glatirameracetat dienen dabei lediglich als Brückenkomparatoren in indirekten Vergleichen bzw. Netzwerken. Ein Wechsel innerhalb der Basistherapeutika der Wirksamkeitskategorie 1 (z. B. von Dimethylfumarat auf Teriflunomid) wird gemäß aktueller AWMF-Leitlinie zwar nicht mehr empfohlen [5], kann allerdings je nach individuellem Krankheitsverlauf und Präferenz der Patientinnen und Patienten eine Therapieoption darstellen und ist auch von der Zulassung der jeweiligen Wirkstoffe umfasst. Unabhängig davon, ob ein solcher Wechsel in der klinischen Praxis üblich ist, ist die Frage nach möglichen Vor- oder Nachteilen für Patientinnen und Patienten weiterhin offen und wird daher in der vorliegenden Bewertung untersucht.

Anforderungen an Studien in der Indikation RRMS

In den letzten Jahren wird zunehmend die Notwendigkeit einer patientenindividuellen Therapie der MS diskutiert [5,50,51]. Neben der Schubreduktion sind dabei auch Schwere der Schübe, Behinderungsprogression, Nebenwirkungen, Symptome der MS und die gesundheitsbezogene Lebensqualität zu berücksichtigen, da sie auch unabhängig voneinander Anlass für Therapieentscheidungen sein können [43].

Behandlungsentscheidungen zwischen Betroffenen und behandelnden Ärzten hängen von einer Vielzahl von Faktoren ab, die neben Wirksamkeit und Sicherheit einzelner Arzneimittel auch die Lebensumstände und persönlichen Präferenzen der Betroffenen umfassen [51]. Studien zu Patientenpräferenzen zeigten außerdem, dass keineswegs nur die Reduktion der Schubrate und die Behinderungsprogression für Betroffene bedeutsam sind, sondern maßgeblich auch Nebenwirkungen, Symptome der Erkrankung und die gesundheitsbezogene Lebensqualität [52-56].

Das Spektrum klinischer Studien sollte die zuvor genannten Anforderungen widerspiegeln, um patientenrelevante Aussagen für den Versorgungsalltag zu ermöglichen. Zu berücksichtigen sind dabei unterschiedliche Therapiestrategien [57]. Randomisierte kontrollierte Studien (RCTs), die eine Eskalationstherapie (z. B. Fingolimod, Ocrelizumab) mit einer Basistherapie (z. B. Interferon, Glatirameracetat) bei starrem Therapieschema vergleichen, erfüllen diese Anforderung allein nicht [58]. Vielmehr wäre anhand von Schubrate, Schwere der Schübe, allgemeiner Symptomatik, gesundheitsbezogener Lebensqualität und Nebenwirkungen für einzelne Patientinnen und Patienten zu prüfen, ob eine dauerhafte Eskalationstherapie überhaupt angebracht ist. Auch in der aktuellen AWMF-Leitlinie zur MS wird für verschiedene Wirkstoffe empfohlen, bei Fehlen von Krankheitsaktivität eine Therapiepause zu erwägen, die Therapie auszusetzen und die Betroffenen über Nutzen und Risiken verschiedener Strategien (Fortführung, Aussetzen, Deeskalation) aufzuklären [5]. Daraus ergeben sich Anforderungen

an Studiendesigns, die nicht nur den Vergleich einzelner Wirkstoffe, sondern den Vergleich ganzer Therapiestrategien ermöglichen (siehe Kapitel 2).

Aufgrund der Vielgestaltigkeit der Symptome und ihrer Wirkung auf den Alltag der Betroffenen kommt patientenberichteten Endpunkten (PROs) eine besondere Bedeutung bei der Nutzenbewertung zu. Gesundheitsbezogene Lebensqualität und Symptomlast müssen gemeinsam betrachtet werden, um Aussagen über den Nutzen von MS-Therapien zu treffen. Hinzu kommt die Schwere der Schübe im Sinne (bleibender) funktioneller Beeinträchtigung, die über die bloße Erfassung der Schubrate nicht abgebildet wird.

Da RRMS eine lebenslange Erkrankung ist, die mitunter erst nach Jahren eine signifikante Behinderung nach sich zieht, sind Beobachtungen über mehrere Jahre notwendig. Dies gilt insbesondere für die Erhebung der Behinderungsprogression. Zwar wird häufig bereits nach 12 bis 24 Wochen eine sogenannte „bestätigte“ Progression festgestellt, allerdings kann eine zu kurze Beobachtungsdauer zu einer Überschätzung des Therapieeffekts führen, weil schubbedingte und progressionsbedingte Behinderung nicht zuverlässig unterschieden werden können [5,59].

Eine Netzwerk-Metaanalyse (NMA), die alle 10 beauftragten Wirkstoffe bei hochaktiver RRMS einbezieht, existiert bislang nicht. Bisherige Arbeiten umfassen nur Teile des Wirkstoffspektrums oder beziehen sich nicht auf die hochaktive RRMS [60-63].

2 Fragestellung

Das Ziel der vorliegenden Untersuchung ist

- die vergleichende Nutzenbewertung der Wirkstoffe Alemtuzumab, Cladribin, Dimethylfumarat, Fingolimod, Natalizumab, Ocrelizumab, Ofatumumab, Ozanimod, Ponesimod und Teriflunomid untereinander

zur Behandlung von erwachsenen Patientinnen und Patienten mit hochaktiver RRMS trotz vollständiger und angemessener Behandlung mit mindestens 1 krankheitsmodifizierenden Therapie. Die Bewertung erfolgt hinsichtlich patientenrelevanter Endpunkte.

Beim Vergleich des Nutzens sind aufgrund der in Kapitel 1 genannten hohen Individualität der Krankheitsverläufe verschiedene Therapiestrategien denkbar:

- Eskalation von einer Basistherapie (etwa Interferone oder Glatirameracetat) auf eine hochwirksamere Therapie (bspw. Fingolimod, Ocrelizumab)
- Deeskalation, also Aussetzen der Therapie oder Wechsel zu einer Basistherapie bei Fehlen von Krankheitsaktivität, nicht tolerierbaren Nebenwirkungen, geplanter Schwangerschaft oder hohem Alter
- Wechsel zu einer anderen Basis- bzw. Eskalationstherapie

Daraus ergeben sich folgende Fragestellungen:

Tabelle 1: Fragestellungen der Nutzenbewertung

| Fragestellung | Vergleich |
|---------------|--|
| 1 | Eskalationstherapie vs. Basistherapie |
| 2 | Eskalationstherapie mit Möglichkeit der Deeskalation vs. Basistherapie |
| 3 | Eskalationstherapie vs. Eskalationstherapie mit Möglichkeit der Deeskalation |
| 4 | Vergleich verschiedener Wirkstoffe innerhalb einer Therapiestrategie |

3 Projektverlauf

3.1 Zeitlicher Verlauf des Projekts

Der Gemeinsame Bundesausschuss (G-BA) hat am 16.07.2020 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung von Alemtuzumab, Cladribin, Dimethylfumarat, Fingolimod, Natalizumab, Ocrelizumab und Teriflunomid zur Behandlung von erwachsenen Patientinnen und Patienten mit hochaktiver RRMS beauftragt.

In die Bearbeitung des Projekts werden externe Sachverständige eingebunden.

Während der Erstellung des Berichtsplans wurden am 24.09.2020 Betroffene zur Diskussion von patientenrelevanten Endpunkten und relevanten Subgruppen konsultiert.

Der Berichtsplan in der Version 1.0 vom 18.12.2020 wurde am 28.12.2020 auf der Website des IQWiG veröffentlicht und zur Anhörung gestellt. Bis zum 26.01.2021 konnten schriftliche Stellungnahmen eingereicht werden. Die Dokumentation der Anhörung zum Berichtsplan ist auf der Website des IQWiG veröffentlicht.

Der Berichtsplan in der Version 2.0 vom 12.08.2021 wurde am 20.08.2021 auf der Webseite des IQWiG veröffentlicht, und beinhaltet die Änderungen, die sich aus der Anhörung ergeben haben.

Der G-BA hat den Auftrag an das IQWiG am 19.11.2021 um die Bewertung der Wirkstoffe Ofatumumab, Ozanimod und Ponesimod zur Behandlung von erwachsenen Patientinnen und Patienten mit hochaktiver RRMS erweitert.

Der vorliegende Berichtsplan in der Version 2.1 beinhaltet die Änderungen, die sich aus der Auftragsenerweiterung um die 3 genannten Wirkstoffe ergeben. Da diese Auftragsenerweiterung keine Auswirkung auf die grundsätzliche Methodik der Nutzenbewertung hat, wird der vorliegende Berichtsplan nicht erneut zur Anhörung gestellt.

Auf Basis des Berichtsplans wird die vorläufige Bewertung vorgenommen. Diese wird in einem Vorbericht veröffentlicht, der zur Anhörung gestellt wird. Der Vorbericht wird gegebenenfalls zusätzlich einem externen Review unterzogen. Im Anschluss an die Anhörung zum Vorbericht erstellt das IQWiG einen Abschlussbericht. Dieser Bericht wird an den G-BA übermittelt und grundsätzlich 4 Wochen später auf der Website des IQWiG veröffentlicht. An selber Stelle wird auch die Dokumentation der Anhörung zum Vorbericht veröffentlicht.

Der Zeitplan für alle Arbeitsschritte der Berichterstellung für dieses Projekt ist auf der Website des IQWiG dargelegt.

3.2 Dokumentation der Änderungen im Projektverlauf

Berichtsplan 2.1 im Vergleich zum Berichtsplan 2.0

- Ergänzung der Wirkstoffe Ofatumumab, Ozanimod und Ponesimod (Kapitel 1 und 2, Abschnitte 4.1.2, 4.1.7, 4.2 und 4.2.1)

Aus der Ergänzung der Wirkstoffe ergeben sich keine weiteren Anpassungen der Methodik. Diese bleibt gegenüber dem Berichtsplan 2.0 somit unverändert.

Berichtsplan 2.0 im Vergleich zum Berichtsplan 1.0

- Erweiterung der Operationalisierung der hochaktiven Erkrankung, um auch Patientinnen und Patienten mit schwerer, d. h. funktionell erheblich beeinträchtigender Schubaktivität, aber ohne auffällige MRT-Befunde sowie Patientinnen und Patienten, bei denen ausschließlich eine hohe MRT-Aktivität vorliegt, in der Bewertung zu berücksichtigen (Kapitel 1 und Abschnitt 4.1.1)
- Spezifizierung der Angaben zur vollständigen und angemessenen Vorbehandlung anhand der Latenzzeiten der spezifischen Wirkstoffe im Zusammenhang mit der Operationalisierung der hochaktiven Erkrankung (Kapitel 1 und Abschnitt 4.1.1)
- Referenzierung der nach Veröffentlichung des Berichtsplans 1.0 erschienenen aktuellen AWMF-Leitlinie zu Diagnostik und Therapie der multiplen Sklerose und entsprechende Anpassung der Neuerungen (Wechsel von Stufentherapieschema auf die Einteilung in Wirksamkeitskategorien) (Kapitel 1)
- Konkretisierungen zu den Fragestellungen der Bewertung (Erläuterung zur Untersuchung des Wechsels innerhalb der Basistherapien als relevante Fragestellung für Patientinnen und Patienten) (Kapitel 1 und 2)
- Aufnahme des Endpunkts schwerwiegende sekundäre Autoimmunerkrankungen (Abschnitt 4.1.3)
- Konkretisierung, dass auch versorgungsnahe, randomisierte Registerstudien für die Bewertung relevant sind (Abschnitt 4.1.4)
- Verlängerung der Studiendauer für den Einschluss in die Bewertung von mindestens 12 Monaten auf mindestens 24 Monate für mindestens einen Teil der Studienpopulation, um dem chronischen Charakter der Erkrankung Rechnung zu tragen und Veränderungen, vor allem der Behinderung der Patientinnen und Patienten, besser erfassen zu können (Abschnitte 4.1.5 und 4.1.7)
- Spezifizierung, dass für binäre Endpunkte auch andere Effektmaße als das relative Risiko herangezogen werden können (Abschnitt 4.3.3)
- Anpassung der Operationalisierung des möglichen Effektmodifikators Schweregrad der Erkrankung: anstelle des Fortschreitens der Erkrankung wird der EDSS-Status für die Unterteilung herangezogen, da für die Einteilung in rasch fortschreitende schwere RRMS

vs. sonstige hochaktive RRMS keine allgemein anerkannte Definition verfügbar ist
(Abschnitt 4.3.5)

- Ergänzung von Subgruppenanalysen nach Definition der hochaktiven Erkrankung, um die Auswirkungen der verschiedenen Definitionen auf die Ergebnisse der Bewertung zu untersuchen (ausschließlich klinisch vs. klinisch mit MRT-Aktivität vs. ausschließlich MRT-basiert) (Abschnitt 4.3.5)

4 Methoden

Diese Bewertung wird auf Grundlage der Allgemeinen Methoden 6.0 [29] erstellt.

4.1 Kriterien für den Einschluss von Studien in die Untersuchung

4.1.1 Population

In die Bewertung werden Studien mit Patientinnen und Patienten mit hochaktiver RRMS aufgenommen, die eine vollständige und angemessene Vorbehandlung mit mindestens 1 krankheitsmodifizierenden Therapie erhalten haben.

Ein hochaktiver Verlauf der RRMS ist dabei definiert als:

- 1) Ausschließlich klinisch: schwere Schübe mit erheblicher funktioneller Beeinträchtigung, unabhängig vom Läsionsgeschehen
- 2) Klinisch in Verbindung mit MRT-Aktivität: Schübe ohne erhebliche funktionelle Beeinträchtigung, im Verbund mit dem Auftreten neuer oder vergrößerter Läsionen
- 3) Ausschließlich MRT-basiert: in Abhängigkeit des Auftretens zahlreicher neuer oder vergrößerter Läsionen

trotz vollständiger und angemessener Vorbehandlung. Zur näheren Spezifizierung der möglichen Definitionen siehe Kapitel 1.

4.1.2 Prüf- und Vergleichsintervention

Alemtuzumab, Cladribin, Dimethylfumarat, Fingolimod, Natalizumab, Ocrelizumab, Ofatumumab, Ozanimod, Ponesimod und Teriflunomid sollen miteinander verglichen werden und sind damit sowohl Prüf- als auch Vergleichsintervention.

Um die Wirkstoffe innerhalb einer gemeinsamen Analyse auch indirekt untereinander vergleichen zu können, werden neben Studien, die 2 der genannten Wirkstoffe direkt untereinander vergleichen, auch Studien in die Nutzenbewertung eingeschlossen, die mindestens 1 dieser Wirkstoffe mit 1 möglichen Brückenkomparator vergleichen. Als mögliche Brückenkomparatoren kommen Brückenkomparatoren 1. Grades infrage. Ein Brückenkomparator 1. Grades ist eine Intervention (inklusive Placebo), für die ein direkter Vergleich sowohl mit einer Prüf- / Vergleichsintervention als auch in einer weiteren Studie mit einer anderen Prüf- / Vergleichsintervention vorliegt. Ein solcher Brückenkomparator leistet damit einen Beitrag zur Bildung eines Netzwerks für einen indirekten Vergleich. Die konkrete Festlegung von Brückenkomparatoren, die für die Nutzenbewertung herangezogen werden, ergibt sich daher anhand des Studienpools der relevanten Studien zu den 10 zu bewertenden Wirkstoffen. Studien, die ausschließlich mögliche Brückenkomparatoren miteinander vergleichen, werden dagegen nicht in die Nutzenbewertung eingeschlossen.

Die Wirkstoffe Alemtuzumab, Cladribin, Dimethylfumarat, Fingolimod, Natalizumab, Ocrelizumab, Ofatumumab, Ozanimod, Ponesimod und Teriflunomid werden unter Berücksichtigung einer jeweiligen Therapiestrategie untereinander in einer NMA verglichen. Auf Grundlage dieser Analyse werden die in Kapitel 2 genannten Fragestellungen einzeln beantwortet. Studien, die mittels der in Tabelle 2 beschriebenen Einschlusskriterien als relevant identifiziert werden, werden anhand der geführten Vergleiche den Fragestellungen zugeordnet.

Die 10 Wirkstoffe müssen gemäß des für Deutschland gültigen Zulassungsstatus eingesetzt werden. Bei Nichtansprechen auf eine vollständige und angemessene Vortherapie ist ein Therapiewechsel bei Studieneinschluss erforderlich (unter Beachtung der Latenzzeit der Wirkstoffe), damit die Intervention nicht mit einer unwirksamen Behandlung verglichen wird (siehe Tabelle 2). Unterschiedliche Applikationsformen und Dosierungen innerhalb eines Wirkstoffs, z. B. Interferon beta-1a subkutan und Interferon beta-1a intramuskulär, werden dabei nicht unterschieden (siehe z. B. [64,65]).

4.1.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevante Endpunkte betrachtet:

- Gesamtmortalität
- bestätigte Krankheitsschübe
 - jährliche Schubrate
 - Patientinnen und Patienten mit bestätigtem Schub
- Behinderung
 - bestätigte Behinderungsprogression (Bestätigung nach mindestens 24 Wochen) anhand der EDSS¹
 - Schweregrad der Behinderung anhand des Multiple-Sclerosis-Functional-Composite-(MSFC)-Index
 - Gehfähigkeit, z. B. anhand des 6-Minute Walk Test
- Fatigue
- Sehstörungen
- gesundheitsbezogene Lebensqualität
- Gesamtrate schwerwiegender unerwünschter Ereignisse (SUEs)
- Gesamtrate der Abbrüche wegen unerwünschter Ereignisse (UEs)
- progressive multifokale Leukenzephalopathie (PML)

¹ Ergebnisse für die EDSS sollten durch Daten anderer Instrumente wie den MSFC ergänzt werden, um ein vollständiges Bild der Behinderungsprogression zu erhalten.

- schwerwiegende Infektionen
- schwerwiegende Neoplasmen
- schwerwiegende sekundäre Autoimmunerkrankungen

Subjektive Endpunkte (zum Beispiel gesundheitsbezogene Lebensqualität) werden nur dann berücksichtigt, wenn sie mit validen Messinstrumenten (zum Beispiel validierten Skalen) erfasst wurden.

4.1.4 Studientypen

RCTs sind, sofern sie methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt wurden, mit der geringsten Ergebnisunsicherheit behaftet. Sie liefern daher die zuverlässigsten Ergebnisse für die Bewertung des Nutzens einer medizinischen Intervention.

Für alle unter 4.1.2 genannten Interventionen und alle unter 4.1.3 genannten Endpunkte ist eine Evaluation im Rahmen von RCTs möglich und praktisch durchführbar.

Für den zu erstellenden Bericht werden daher RCTs als relevante wissenschaftliche Literatur in die Nutzenbewertung einfließen. Darunter fallen auch versorgungsnahe, randomisierte Studien aus Registern.

4.1.5 Studiendauer

Eingeschlossen werden Studien, in denen mindestens ein Teil der Studienpopulation für mindestens 24 Monate beobachtet wurde.

4.1.6 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfasst sein.

4.1.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

Tabelle 2: Übersicht über die Kriterien für den Studieneinschluss

| Einschlusskriterien | |
|---|---|
| E1 | erwachsene Patientinnen und Patienten mit hochaktiver RRMS trotz angemessener und vollständiger Vorbehandlung mit einer krankheitsmodifizierenden Therapie (siehe auch Abschnitt 4.1.1) |
| E2 | Prüfintervention: Behandlung mit Alemtuzumab, Cladribin, Dimethylfumarat, Fingolimod, Natalizumab, Ocrelizumab, Ofatumumab, Ozanimod, Ponesimod oder Teriflunomid gemäß Zulassung in mindestens 1 Studienarm (siehe auch Abschnitt 4.1.2) |
| E3 | Vergleichsintervention: jegliche andere medikamentöse Therapie einschließlich der unter E2 genannten Prüfinterventionen oder Placebo (siehe auch Abschnitt 4.1.2) |
| E4 | patientenrelevante Endpunkte wie in Abschnitt 4.1.3 formuliert |
| E5 | RCT |
| E6 | Studiendauer von mindestens 24 Monaten für mindestens einen Teil der Studienpopulation |
| E7 | Publikationssprache: Deutsch oder Englisch |
| E8 | Vollpublikation verfügbar ^a |
| <p>a. Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [66] oder ein Bericht über die Studie, der den Kriterien des CONSORT-Statements [67] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zur Studienmethodik und zu den Studienergebnissen nicht vertraulich sind.</p> <p>CONSORT: Consolidated Standards of Reporting Trials; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; RCT: randomisierte kontrollierte Studie; RRMS: schubförmig remittierende multiple Sklerose</p> | |

4.1.8 Vorgehen im Falle einer Zulassungsänderung im Projektverlauf

Sofern sich im Projektverlauf Änderungen im Zulassungsstatus der zu bewertenden Interventionen ergeben, werden die Kriterien für den Studieneinschluss gegebenenfalls an die neuen Zulassungsbedingungen angepasst. Die jeweils vorgenommenen Änderungen werden im Vorbericht beziehungsweise im Abschlussbericht explizit vermerkt.

4.1.9 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen

Für die Einschlusskriterien E1 (Population), E2 (Prüfintervention, bezogen auf die Interventionsgruppe der Studie) und E3 (Vergleichsintervention, bezogen auf die Vergleichsgruppe der Studie) reicht es aus, wenn bei mindestens 80 % der eingeschlossenen Patientinnen und Patienten diese Kriterien erfüllt sind. Liegen für solche Studien Subgruppenanalysen für Patientinnen und Patienten vor, die die Einschlusskriterien erfüllen, wird auf diese Analysen zurückgegriffen. Studien, bei denen die Einschlusskriterien E1, E2 und E3 bei weniger als 80 % erfüllt sind, werden nur dann eingeschlossen, wenn Subgruppenanalysen für Patientinnen und Patienten vorliegen, die die Einschlusskriterien erfüllen.

4.2 Informationsbeschaffung

Die Informationsbeschaffung beruht maßgeblich auf Herstelleranfragen zu den Wirkstoffen Alemtuzumab, Cladribin, Dimethylfumarat, Fingolimod, Natalizumab, Ocrelizumab, Ofatumumab, Ozanimod, Ponesimod und Teriflunomid. Suchen in Studienregistern und bibliografischen Datenbanken dienen dazu, die Angaben der Hersteller auf Vollständigkeit zu

überprüfen und relevante Studien von Dritten zu identifizieren. Es ist nicht das Ziel der Informationsbeschaffung, alle öffentlich verfügbaren Dokumente zu einer relevanten Studie zu identifizieren.

4.2.1 Informationsquellen

Primäre Informationsquelle

- Herstelleranfragen

Die Herstelleranfragen ermöglichen den Überblick über alle von Herstellern durchgeführten Studien unabhängig vom Publikationsstatus. Für potenziell relevante Studien aus dieser Übersicht werden vollständige Studienunterlagen (i. d. R. vollständige Studienberichte) angefordert und in die Bewertung einbezogen. Sofern erforderlich werden Zusatzanalysen angefordert. Alle Informationen zur Methodik und zu Ergebnissen, die in die Nutzenbewertung eingehen, werden im Bericht des IQWiG veröffentlicht. Angefragt werden folgende Hersteller:

- Biogen GmbH (Dimethylfumarat, Natalizumab)
- Bristol Myers Squibb GmbH & Co. KGaA (Ozanimod)
- Janssen-Cilag GmbH (Ponesimod)
- Merck Serono GmbH (Cladribin)
- Novartis Pharma GmbH (Fingolimod, Ofatumumab)
- Roche Pharma AG (Ocrelizumab)
- Sanofi-Aventis Deutschland GmbH (Alemtuzumab, Teriflunomid)

Generikahersteller werden in diesem Kontext nicht berücksichtigt, da regelhaft ausschließlich Äquivalenzstudien zu erwarten sind.

Informationsquellen zur Identifikation weiterer Studien

Die Überprüfung der Vollständigkeit der Herstellerunterlagen erfolgt vorrangig anhand einer Suche in Studienregistern, da Studien neuerer Arzneimittel nahezu vollständig registriert sind.

Für ältere Studien (vor 2005) und Studien Dritter ohne Herstellerbeteiligung trifft dies jedoch nur eingeschränkt zu. Daher wird für alle Wirkstoffe zusätzlich zur Suche in Studienregistern eine fokussierte Recherche in MEDLINE durchgeführt.

Zur Überprüfung der Vollständigkeit der Herstellerunterlagen und Identifikation weiterer relevanter Studien werden daher folgende Informationsquellen durchsucht:

- Studienregister
 - U.S. National Institutes of Health. ClinicalTrials.gov

- World Health Organization. International Clinical Trials Registry Platform Search Portal
- European Medicines Agency. EU Clinical Trials Register
- bibliografische Datenbanken
 - MEDLINE
- G-BA-Website und IQWiG-Website

Darüber hinaus werden Stellungnahmen auf Informationen zu relevanten Studien gesichtet. Gegebenenfalls können auch Anfragen bei Studienautoren direkt erfolgen.

4.2.2 Selektion relevanter Studien

Selektion relevanter Studien bzw. Dokumente aus den Ergebnissen der bibliografischen Recherche

Die in bibliografischen Datenbanken identifizierten Treffer werden in einem 1. Schritt anhand ihres Titels und, sofern vorhanden, Abstracts in Bezug auf ihre potenzielle Relevanz bezüglich der Einschlusskriterien (siehe Tabelle 2) bewertet. Als potenziell relevant erachtete Dokumente werden in einem 2. Schritt anhand ihres Volltextes auf Relevanz geprüft. Beide Schritte erfolgen durch 2 Personen unabhängig voneinander. Diskrepanzen werden durch Diskussion zwischen den beiden aufgelöst.

Alle Bewertungen erfolgen durch 2 Personen unabhängig voneinander. Diskrepanzen werden durch Diskussion zwischen den beiden aufgelöst.

Selektion relevanter Studien bzw. Dokumente aus weiteren Informationsquellen

Rechercheergebnisse aus Studienregistern werden von 2 Personen unabhängig voneinander in Bezug auf ihre Relevanz bewertet. Rechercheergebnisse aus den darüber hinaus berücksichtigten Informationsquellen werden von 1 Person in Bezug auf Studien gesichtet. Die identifizierten Studien werden dann auf ihre Relevanz geprüft. Der gesamte Prozess wird anschließend von einer 2. Person überprüft. Sofern in einem der genannten Selektionsschritte Diskrepanzen auftreten, werden diese jeweils durch Diskussion zwischen den beiden Personen aufgelöst.

4.3 Informationsbewertung und Synthese

4.3.1 Darstellung der Einzelstudien

Alle für die Nutzenbewertung notwendigen Informationen werden aus den Unterlagen zu den eingeschlossenen Studien in standardisierte Tabellen extrahiert. Studien, deren Vergleichsintervention keinen geeigneten Brückenkomparator 1. Grades darstellt, werden nicht in die weiteren Prüf- und Analyseschritte einbezogen. Für solche Studien werden ausschließlich Studien- und Interventionscharakteristika dargestellt.

Die Ergebnisse zu den in den Studien berichteten patientenrelevanten Endpunkten werden im Bericht vergleichend beschrieben.

Die relevanten Ergebnisse werden endpunktspezifisch pro Studie auf ihr jeweiliges Verzerrungspotenzial überprüft (siehe Abschnitt 4.3.2). Anschließend werden die Informationen zusammengeführt und analysiert. Wenn möglich werden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die unter 4.3.3 bis 4.3.5 beschriebenen Verfahren eingesetzt. Eine abschließende zusammenfassende Bewertung der Informationen erfolgt in jedem Fall.

Ergebnisse fließen in der Regel nicht in die Nutzenbewertung ein, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Patientinnen und Patienten basieren, das heißt, wenn der Anteil der Patientinnen und Patienten, die nicht in der Auswertung berücksichtigt werden, größer als 30 % ist.

Die Ergebnisse werden auch dann nicht in die Nutzenbewertung einbezogen, wenn der Unterschied der Anteile nicht berücksichtigter Patientinnen und Patienten zwischen den Gruppen größer als 15 Prozentpunkte ist.

4.3.2 Bewertung des Verzerrungspotenzials der Ergebnisse

Das Verzerrungspotenzial der Ergebnisse wird endpunktspezifisch für jede in die Nutzenbewertung eingeschlossene Studie bewertet. Für Studien, die ausschließlich einen indirekten Vergleich von Prüfinerventionen ermöglichen, erfolgt die Bewertung des Verzerrungspotenzials nur dann, wenn es als Faktor bei der Überprüfung der Strukturqualität zu untersuchen oder für die Ableitung der Beleglage ausschlaggebend ist. Ausschlaggebend für die Beleglage ist das Verzerrungspotenzial bei einem statistisch signifikanten Unterschied auf Basis eines indirekten Vergleichs, bei dem für mindestens 1 Prüfinervention nur 1 Studie vorliegt.

Zur Bewertung des Verzerrungspotenzials werden insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Kriterien systematisch extrahiert und bewertet:

A: Kriterien zur endpunktübergreifenden Bewertung des Verzerrungspotenzials der Ergebnisse

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung der Patientin oder des Patienten sowie der behandelnden Personen
- ergebnisunabhängige Berichterstattung

B: Kriterien zur endpunktspezifischen Bewertung des Verzerrungspotenzials der Ergebnisse

- Verblindung der Endpunkterheber

- Umsetzung des Intention-to-treat(ITT)-Prinzips
- ergebnisunabhängige Berichterstattung

Für die Ergebnisse randomisierter Studien wird das Verzerrungspotenzial zusammenfassend als niedrig oder hoch eingestuft. Wird bereits hinsichtlich der unter (A) aufgeführten Kriterien ein endpunktübergreifend hohes Verzerrungspotenzial festgestellt, gilt dieses damit für alle Ergebnisse aller Endpunkte als hoch, unabhängig von der Bewertung endpunktspezifischer Aspekte. Andernfalls finden anschließend die unter (B) genannten Kriterien pro Endpunkt Berücksichtigung.

4.3.3 Netzwerk-Metaanalysen

Um die Prüfinterventionen innerhalb einer gemeinsamen Analyse vergleichen zu können, werden bevorzugt NMAs eingesetzt. Dadurch werden auch solche Vergleiche ermöglicht, für die keine direkt vergleichende Evidenz vorliegt. Ist die Durchführung einer NMA nicht möglich, beruht die Analyse ausschließlich, falls vorhanden, auf direkten Vergleichen von 2 der genannten Prüfinterventionen. Sollte keine NMA möglich sein und sollten keine direkten Vergleiche von Prüfinterventionen vorliegen, werden paarweise adjustierte indirekte Vergleiche nach der Methode von Bucher [68] durchgeführt. Für die statistische Auswertung werden die Ergebnisse aus ITT-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet.

Die Voraussetzung dafür, eine NMA oder indirekte Vergleiche durchzuführen und zu interpretieren, ist eine ausreichende Strukturqualität. Diese liegt vor, wenn jeweils die Annahme von Ähnlichkeit, Homogenität und Konsistenz erfüllt ist:

1) Überprüfung der Ähnlichkeitsannahme

Ähnlichkeit liegt vor, wenn Faktoren mit möglichem Einfluss auf die Effektschätzungen in allen Studien vergleichbar ausgeprägt sind. Dazu zählen:

- Studiencharakteristika (z. B. Studiendauer)
- Patientencharakteristika (z. B. Alter, Geschlecht)
- Krankheitscharakteristika (z. B. Schubrate und EDSS-Score, jeweils zu Studienbeginn)
- Interventionscharakteristika (z. B. Begleittherapien)
- Endpunktoperationalisierung (z. B. Effektmaß)

Sind Faktoren, die auf Ähnlichkeit geprüft werden, in einem bedeutsamen Ausmaß unterschiedlich ausgeprägt, wird nicht mehr von einer ausreichenden Ähnlichkeit ausgegangen und es erfolgt keine gemeinsame Analyse. Stattdessen werden für Teilmengen von Studien, in denen vergleichbare Ausprägungen vorliegen, jeweils

Analysen durchgeführt. Beispielsweise können Teilnetzwerke oder nur einzelne Vergleiche betrachtet werden.

2) Überprüfung der Homogenitätsannahme

Homogenität liegt vor, wenn für jeden Vergleich die Ergebnisse der direkt vergleichenden Studien (Vergleich einer Prüfintervention mit 1 oder mehreren anderen Prüfinerventionen oder mit einem Brückenkomparator) homogen sind. Von homogenen Effekten wird ausgegangen, wenn durch den entsprechenden statistischen Test [69] keine bedeutsame Heterogenität angezeigt wird. Falls heterogene Effekte vorliegen, wird untersucht, welche Faktoren (unter anderem klinische und methodische) diese Heterogenität möglicherweise erklären können (siehe Aufzählungspunkt 4). Es können außerdem weitere Ergebnisse aus Regressions- und Subgruppenanalysen einbezogen werden, wenn entsprechende Daten verfügbar sind. Gegebenenfalls erfolgen eigene Analysen in Form von Metaregressionen oder Metaanalysen unter Kategorisierung der Studien bezüglich der möglichen Einflussfaktoren.

3) Überprüfung der Konsistenzannahme

Für valide Ergebnisse ist die Konsistenz innerhalb des Netzwerks notwendig. Diese liegt vor, wenn die Schätzungen aus direkten und indirekten Vergleichen übereinstimmen. Die Konsistenz wird mit der Methode der Knotenteilung (Node Splitting [70]) überprüft: Wenn für einen Vergleich von 2 Prüfinerventionen direkte Evidenz vorliegt, der diesbezügliche Effekt aber auch indirekt geschätzt werden kann, dann werden 2 Effektschätzungen betrachtet: Die Effektschätzung aus direkter Evidenz wird mit der Effektschätzung verglichen, die sich aus der NMA unter Ausschluss dieser direkten Evidenz ergibt. Falls ein Test auf Gleichheit beider Effekte keinen statistisch signifikanten Unterschied (Niveau $\alpha = 0,05$) nachweist, wird davon ausgegangen, dass an der betrachteten Stelle des Netzwerks keine Inkonsistenz vorliegt.

Werden Unterschiede festgestellt, wird die Ähnlichkeit der Studien nochmals überprüft und das Netzwerk gegebenenfalls angepasst (siehe Aufzählungspunkt 4).

4) Vorgehen bei bedeutsamer Heterogenität oder Inkonsistenz

Liegt bedeutsame Heterogenität oder Inkonsistenz vor, wird in den ersten Schritten untersucht, ob Faktoren, für die bei Überprüfung der Ähnlichkeit der Studien Unsicherheiten entstanden, erklärende Faktoren für bedeutsame Heterogenität oder Inkonsistenz sein können. Diese Faktoren werden nacheinander untersucht (zum Beispiel Alter, den Krankheitsverlauf beeinflussende Faktoren). Sind diese Faktoren keine mögliche Erklärung für bedeutsame Heterogenität oder Inkonsistenz, werden nachfolgend weitere Faktoren separat und nacheinander untersucht, die im Berichtsplan für die Prüfung der Ähnlichkeit der Studien spezifiziert waren. Ist keiner der Faktoren, die für die Prüfung der Ähnlichkeit der Studien berücksichtigt wurden, eine mögliche Erklärung für bedeutsame Heterogenität oder Inkonsistenz, wird das Verzerrungspotenzial als möglicher erklärender Faktor untersucht. Wird mindestens 1 Studie in einem Prüfschritt

ausgeschlossen, wird für den so reduzierten Studienpool erneut die Homogenitäts- und Konsistenzannahme (sofern auf Basis der Studienlage möglich) geprüft.

NMAs und indirekte Vergleiche werden zwar auch durchgeführt, wenn nicht alle Annahmen der Strukturqualität überprüfbar sind, weil beispielsweise keine ausreichende Anzahl von Studien vorliegt. In keinem Fall sollte aber eine der Annahmen offensichtlich verletzt sein.

NMAs werden vorzugsweise mit frequentistischen Methoden (Modelle mit zufälligen Effekten) berechnet [71]. Bayessche Verfahren werden ersatzweise verwendet. Die Auswahl der Modelle für die jeweiligen Metaanalysen erfolgt gemäß den Kriterien, die in den Allgemeinen Methoden 6.0 [58] genannt sind. Als Ergebnisse der jeweiligen Analysen werden die resultierenden gemeinsamen (gepoolten) geschätzten Effekte inklusive Konfidenzintervallen dargestellt.

Als Effektmaß wird für stetige Variablen die Mittelwertdifferenz herangezogen. Wenn aus der finalen NMA mindestens 1 statistisch signifikanter Effekt resultiert und keine validierte bzw. etablierte skalenspezifische Irrelevanzschwelle vorliegt, wird in einer anschließenden Analyse als Effektmaß die standardisierte Mittelwertdifferenz (SMD, in Form von Hedges' g) verwendet. In diese anschließende Analyse werden dieselben Studien der finalen NMA einbezogen. Dabei wird von einem relevanten Effekt ausgegangen, wenn das 95 %-Konfidenzintervall (KI) eines Effekts vollständig unterhalb von $-0,2$ beziehungsweise vollständig oberhalb von $0,2$ liegt.

Bei binären Variablen werden Analysen primär anhand des relativen Risikos durchgeführt. In begründeten Fällen kommen auch andere Effektmaße zum Einsatz. Bei kategorialen Variablen wird ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und von den verfügbaren Daten verwendet [72]. Falls die für eine Analyse notwendigen Schätzungen für Lage und Streuung in den Studienunterlagen nicht vorliegen, werden diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt.

In der Regel werden für die vorliegende Nutzenbewertung ausreichend ähnliche Operationalisierungen, Analysearten und Zeitpunkte berücksichtigt, die mehrheitlich in den eingeschlossenen Studien vorliegen.

Im Kontext der NMAs werden folgende Begriffe verwendet:

- Vorläufige Analyse: 1. Analyse auf Basis eines Studienpools, der aus der Ähnlichkeitsprüfung hervorgegangen ist und für den keine bedeutsame Heterogenität oder Inkonsistenz beobachtet wird (sofern die Prüfungen möglich sind). An die vorläufige Analyse können sich Sensitivitätsanalysen anschließen (Abschnitt 4.3.4).
- NMA: finale Analyse für einen Endpunkt, auf der die Ableitung der Beleglage basiert

Forest Plots werden für die vorläufige Analyse dargestellt, allerdings ausschließlich bei bedeutsamer Heterogenität. Der verwendete Programmcode wird im Vorbericht veröffentlicht.

4.3.4 Sensitivitätsanalysen

Bestehen Zweifel an der Robustheit von Ergebnissen wegen methodischer Faktoren, die beispielsweise durch die Wahl bestimmter Cut-off-Werte, Ersetzungsstrategien für fehlende Werte, Erhebungszeitpunkte oder Effektmaße begründet sein können, ist geplant, den Einfluss solcher Faktoren in Sensitivitätsanalysen zu untersuchen. Das Ergebnis solcher Sensitivitätsanalysen kann die Sicherheit der aus den beobachteten Effekten abgeleiteten Aussagen beeinflussen. Ein als nicht robust eingestufte Effekt kann zum Beispiel dazu führen, dass nur ein Hinweis auf anstelle eines Belegs für einen (höheren) Nutzen attestiert wird (zur Ableitung von Aussagen zur Beleglage siehe Abschnitt 4.3.6).

4.3.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse werden hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht. Ziel ist es, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Für einen Nachweis unterschiedlicher Effekte ist die auf einem Homogenitäts- beziehungsweise Interaktionstest basierende statistische Signifikanz Voraussetzung. In die Untersuchung werden die vorliegenden Ergebnisse aus Regressionsanalysen, die Interaktionsterme beinhalten, und aus Subgruppenanalysen einbezogen. Außerdem erfolgen eigene Analysen in Form von Metaregressionen oder Metaanalysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren. Subgruppenanalysen werden nur durchgeführt, falls jede Subgruppe mindestens 10 Personen umfasst und bei binären Daten mindestens 10 Ereignisse in 1 der Subgruppen aufgetreten sind. Es ist vorgesehen, folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einzubeziehen:

- Geschlecht
- Alter
- Schweregrad der Erkrankung (anhand des EDSS-Status)
- Definition hochaktiver Erkrankung (ausschließlich klinisch vs. klinisch mit MRT-Aktivität vs. ausschließlich MRT-basiert)
- geografische Region

Sollten sich aus den verfügbaren Informationen weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden.

Bei Identifizierung möglicher Effektmodifikatoren erfolgt gegebenenfalls eine Präzisierung der aus den beobachteten Effekten abgeleiteten Aussagen. Beispielsweise kann der Beleg eines (höheren) Nutzens auf eine spezielle Subgruppe von Patientinnen und Patienten eingeschränkt werden (zur Ableitung von Aussagen zur Beleglage siehe Abschnitt 4.3.6).

4.3.6 Aussagen zur Beleglage

Für jeden Endpunkt wird eine Aussage zur Beleglage des (höheren) Nutzens und (höheren) Schadens getroffen. Dabei sind 4 Abstufungen der Aussagesicherheit möglich: Es liegt entweder ein Beleg (höchste Aussagesicherheit), ein Hinweis (mittlere Aussagesicherheit), ein Anhaltspunkt (schwächste Aussagesicherheit) oder keine dieser 3 Situationen vor. Der letzte Fall tritt ein, wenn keine Daten vorliegen oder die vorliegenden Daten keine der 3 übrigen Aussagen zulassen. In diesem Fall wird die Aussage „Es liegt kein Anhaltspunkt für einen (höheren) Nutzen oder (höheren) Schaden vor“ getroffen.

Die Aussagesicherheit für den Nachweis eines höheren Nutzens oder höheren Schadens wird bei einer NMA für jeden Vergleich der genannten Wirkstoffe unter Berücksichtigung der Ergebnissicherheit der direkten und der indirekten Evidenz bewertet.

Die Aussagesicherheit eines direkten Vergleichs richtet sich nach der Anzahl verfügbarer Studien, der qualitativen und quantitativen Sicherheit ihrer Ergebnisse sowie der Homogenität der Ergebnisse bei mehreren Studien. Die qualitative Ergebnissicherheit ist abhängig vom Design der Studie. Ergebnisse randomisierter Studien mit niedrigem Verzerrungspotenzial haben eine hohe, Ergebnisse randomisierter Studien mit hohem Verzerrungspotenzial eine mäßige qualitative Ergebnissicherheit.

Die Aussagesicherheit eines Vergleichs aus einer NMA hängt neben den vorhergehend beschriebenen Aspekten von der Strukturqualität ab. Indirekte Vergleiche, die ausschließlich auf Basis indirekter Evidenz erfolgen, weisen bei guter Strukturqualität und einer ausreichenden Zahl von Studien mit niedrigem Verzerrungspotenzial höchstens eine mäßige Aussagesicherheit auf. Sind einzelne Aspekte der Strukturqualität nicht bewertbar, ist die Ergebnissicherheit eines indirekten Vergleichs höchstens gering.

Abschließend erfolgt eine endpunktübergreifende Bewertung des Nutzens. In diese übergreifende Bewertung wird auch die Datenvollständigkeit und die sich daraus möglicherweise ergebende Verzerrung aufgrund von Publikationsbias einbezogen, die zur Einschränkung des Fazits führen kann.

5 Literatur

1. Holstiege J, Steffen A, Goffrier B et al. Epidemiologie der Multiplen Sklerose – eine populationsbasierte deutschlandweite Studie [online]. 2018 [Zugriff: 11.11.2020]. URL: https://www.versorgungsatlas.de/fileadmin/ziva_docs/86/VA-86-Multiple%20Sklerose-Bericht-V13_Cor..pdf.
2. Kip M, Schönfelder T, Bleß HH. Weißbuch Multiple Sklerose: Versorgungssituation in Deutschland [online]. 2016 [Zugriff: 30.11.2020]. URL: <https://link.springer.com/content/pdf/10.1007%2F978-3-662-49204-8.pdf>.
3. Flachenecker P, Stuke K, Elias W et al. Multiple sclerosis registry in Germany: results of the extension phase 2005/2006. *Dtsch Arztebl Int* 2008; 105(7): 113-119. <https://dx.doi.org/10.3238/arztebl.2008.0113>.
4. Kister I, Bacon TE, Chamot E et al. Natural history of multiple sclerosis symptoms. *Int J MS Care* 2013; 15(3): 146-158. <https://dx.doi.org/10.7224/1537-2073.2012-053>.
5. Deutsche Gesellschaft für Neurologie. Diagnose und Therapie der Multiplen Sklerose, Neuromyelitis-optica-Spektrum-Erkrankungen und MOG-IgG-assoziierten Erkrankungen [online]. 2021 [Zugriff: 25.06.2021]. URL: https://www.awmf.org/uploads/tx_szleitlinien/030-050LG1_S2e_Diagnose-Therapie-Multiplen-Sklerose-Neuromyelitis-Optica-Spektrum-MOG-IgG-assoziierte_Erkrankungen_2021-05_1.pdf.
6. Thompson AJ, Banwell BL, Barkhof F et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol* 2018; 17(2): 162-173. [https://dx.doi.org/10.1016/S1474-4422\(17\)30470-2](https://dx.doi.org/10.1016/S1474-4422(17)30470-2).
7. Kappos L, Wolinsky JS, Giovannoni G et al. Contribution of Relapse-Independent Progression vs Relapse-Associated Worsening to Overall Confirmed Disability Accumulation in Typical Relapsing Multiple Sclerosis in a Pooled Analysis of 2 Randomized Clinical Trials. *JAMA Neurol* 2020; 77(9): 1132-1140. <https://dx.doi.org/10.1001/jamaneurol.2020.1568>.
8. Thompson AJ, Baranzini SE, Geurts J et al. Multiple sclerosis. *Lancet* 2018; 391(10130): 1622-1636. [https://dx.doi.org/10.1016/S0140-6736\(18\)30481-1](https://dx.doi.org/10.1016/S0140-6736(18)30481-1).
9. Lublin FD. New multiple sclerosis phenotypic classification. *Eur Neurol* 2014; 72(Suppl 1): 1-5. <https://dx.doi.org/10.1159/000367614>.
10. Krause I, Kern S, Horntrich A et al. Employment status in multiple sclerosis: impact of disease-specific and non-disease-specific factors. *Mult Scler* 2013; 19(13): 1792-1799. <https://dx.doi.org/10.1177/1352458513485655>.
11. Pucak ML, Carroll KA, Kerr DA et al. Neuropsychiatric manifestations of depression in multiple sclerosis: neuroinflammatory, neuroendocrine, and neurotrophic mechanisms in the pathogenesis of immune-mediated depression. *Dialogues Clin Neurosci* 2007; 9(2): 125-139.

12. Vattakatuchery JJ, Rickards H, Cavanna AE. Pathogenic mechanisms of depression in multiple sclerosis. *J Neuropsychiatry Clin Neurosci* 2011; 23(3): 261-276.
<https://dx.doi.org/10.1176/jnp.23.3.jnp261>.
13. Heesen C, Haase R, Melzig S et al. Perceptions on the value of bodily functions in multiple sclerosis. *Acta Neurol Scand* 2018; 137(3): 356-362.
<https://dx.doi.org/10.1111/ane.12881>.
14. Diaz C, Zarco LA, Rivera DM. Highly active multiple sclerosis: An update. *Mult Scler Relat Disord* 2019; 30: 215-224. <https://dx.doi.org/10.1016/j.msard.2019.01.039>.
15. Ellenberger D, Flachenecker P, Fneish F et al. Aggressive multiple sclerosis: a matter of measurement and timing. *Brain* 2020; 143(11): e97.
<https://dx.doi.org/10.1093/brain/awaa306>.
16. Jacobaeus E, Arrambide G, Pia Amato M et al. Aggressive multiple sclerosis (1): Towards a definition of the phenotype. *Mult Scler* 2020; 26(9): 1031-1044.
<https://dx.doi.org/10.1177/1352458520925369>.
17. Arrambide G, Jacobaeus E, Amato MP et al. Aggressive multiple sclerosis (2): Treatment. *Mult Scler* 2020; 26: 1045-1063. <https://dx.doi.org/10.1177/1352458520924595>.
18. Sanofi Genzyme. LEMTRADA 12 mg Konzentrat zur Herstellung einer Infusionslösung [online]. 2021 [Zugriff: 20.10.2021]. URL: <https://www.fachinfo.de/>.
19. Merck. MAVENCLAD 10 mg Tabletten [online]. 2021 [Zugriff: 02.07.2021]. URL: <https://www.fachinfo.de/>.
20. Biogen. Tecfidera 120 mg/240 mg magensaftresistente Hartkapseln [online]. 2020 [Zugriff: 02.07.2021]. URL: <https://www.fachinfo.de/>.
21. Novartis Pharma. Gilenya 0,25 mg/ - 0,5 mg Hartkapseln [online]. 2021 [Zugriff: 28.10.2021]. URL: <https://www.fachinfo.de/>.
22. Biogen. Tysabri 300 mg [online]. 2021 [Zugriff: 28.10.2021]. URL: <https://www.fachinfo.de/>.
23. Biogen. Tysabri 150 mg Injektionslösung in einer Fertigspritze [online]. 2021 [Zugriff: 28.10.2021]. URL: <https://www.fachinfo.de/>.
24. Roche. Ocrevus [online]. 2021 [Zugriff: 02.07.2021]. URL: <https://www.fachinfo.de/>.
25. Novartis Pharma. Kesimpta 20 mg Injektionslösung in einer Fertigspritze; Kesimpta 20 mg Injektionslösung im Fertigpen [online]. 2021 [Zugriff: 19.10.2021]. URL: <https://www.fachinfo.de/>.
26. Bristol Myers Squibb. Zeposia Hartkapseln [online]. 2021 [Zugriff: 24.11.2021]. URL: <https://www.fachinfo.de/>.
27. Janssen. Ponvory Filmtabletten [online]. 2021 [Zugriff: 19.10.2021]. URL: <https://www.fachinfo.de/>.

28. Sanofi Genzyme. AUBAGIO 7 mg Filmtabletten; AUBAGIO 14 mg Filmtabletten [online]. 2021 [Zugriff: 28.10.2021]. URL: <https://www.fachinfo.de/>.
29. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden; Version 6.0 [online]. 2020 [Zugriff: 30.11.2020]. URL: https://www.iqwig.de/download/Allgemeine-Methoden_Version-6-0.pdf.
30. Ohlmeier C, Gothe H, Haas J et al. Epidemiology, characteristics and treatment of patients with relapsing remitting multiple sclerosis and incidence of high disease activity: Real world evidence based on German claims data. PLoS One 2020; 15(5): e0231846. <https://dx.doi.org/10.1371/journal.pone.0231846>.
31. Spelman T, Freilich J, Anell B et al. Patients With High-disease-activity Relapsing-Remitting Multiple Sclerosis in Real-world Clinical Practice: A Population-based Study in Sweden. Clin Ther 2020; 42(2): 240-250. <https://dx.doi.org/10.1016/j.clinthera.2019.11.018>.
32. NHS England. Treatment Algorithm for Multiple Sclerosis Disease-Modifying Therapies [online]. 2019 [Zugriff: 30.11.2020]. URL: <https://www.england.nhs.uk/commissioning/wp-content/uploads/sites/12/2019/03/Treatment-Algorithm-for-Multiple-Sclerosis-Disease-Modifying-Therapies-08-03-2019-1.pdf>.
33. Gasperini C, Prosperini L, Tintore M et al. Unraveling treatment response in multiple sclerosis: A clinical and MRI challenge. Neurology 2019; 92(4): 180-192. <https://dx.doi.org/10.1212/WNL.00000000000006810>.
34. Rio J, Rovira A, Tintore M et al. Disability progression markers over 6-12 years in interferon-beta-treated multiple sclerosis patients. Mult Scler 2018; 24(3): 322-330. <https://dx.doi.org/10.1177/1352458517698052>.
35. Rio J, Comabella M, Montalban X. Predicting responders to therapies for multiple sclerosis. Nat Rev Neurol 2009; 5(10): 553-560. <https://dx.doi.org/10.1038/nrneurol.2009.139>.
36. Gulani V, Calamante F, Shellock FG et al. Gadolinium deposition in the brain: summary of evidence and recommendations. Lancet Neurol 2017; 16(7): 564-570. [https://dx.doi.org/10.1016/S1474-4422\(17\)30158-8](https://dx.doi.org/10.1016/S1474-4422(17)30158-8).
37. Deutsche Röntgengesellschaft. Gadolinium-haltige Kontrastmittel in der Magnetresonanztomographie [online]. 2019 [Zugriff: 17.12.2020]. URL: <https://www.drg.de/de-DE/3994/mrt-kt/>.
38. Sormani MP, Gasperini C, Romeo M et al. Assessing response to interferon-beta in a multicenter dataset of patients with MS. Neurology 2016; 87(2): 134-140. <https://dx.doi.org/10.1212/WNL.00000000000002830>.
39. Sormani MP, Truffinet P, Thangavelu K et al. Predicting long-term disability outcomes in patients with MS treated with teriflunomide in TEMSO. Neurol Neuroimmunol Neuroinflamm 2017; 4(5): e379. <https://dx.doi.org/10.1212/NXI.0000000000000379>.

40. Saslow L, Li DKB, Halper J et al. An International Standardized Magnetic Resonance Imaging Protocol for Diagnosis and Follow-up of Patients with Multiple Sclerosis: Advocacy, Dissemination, and Implementation Strategies. *Int J MS Care* 2020; 22(5): 226-232. <https://dx.doi.org/10.7224/1537-2073.2020-094>.
41. European Medicines Agency. Tysabri; EPAR; Scientific Discussion [online]. 2007 [Zugriff: 15.07.2021]. URL: https://www.ema.europa.eu/documents/scientific-discussion/tysabri-epar-scientific-discussion_en.pdf.
42. Wattjes MP, Ciccarelli O, Reich DS et al. 2021 MAGNIMS-CMSC-NAIMS consensus recommendations on the use of MRI in patients with multiple sclerosis. *Lancet Neurol* 2021. [https://dx.doi.org/10.1016/S1474-4422\(21\)00095-8](https://dx.doi.org/10.1016/S1474-4422(21)00095-8).
43. Freedman MS, Devonshire V, Duquette P et al. Treatment Optimization in Multiple Sclerosis: Canadian MS Working Group Recommendations. *Can J Neurol Sci* 2020; 47(4): 437-455. <https://dx.doi.org/10.1017/cjn.2020.66>.
44. Lublin FD, Reingold SC, Cohen JA et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology* 2014; 83(3): 278-286. <https://dx.doi.org/10.1212/WNL.0000000000000560>.
45. European Medicines Agency. Guideline on clinical investigation of medicinal products for the treatment of Multiple Sclerosis [online]. 2015 [Zugriff: 10.08.2020]. URL: https://www.ema.europa.eu/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-treatment-multiple-sclerosis_en-0.pdf.
46. Krankheitsbezogenes Kompetenznetz Multiple Sklerose. Qualitätshandbuch MS / NMOSD; Empfehlungen zur Therapie der Multiplen Sklerose / Neuromyelitis-optica-Spektrum-Erkrankungen für Ärzte [online]. 2021 [Zugriff: 28.10.2021]. URL: https://www.kompetenznetz-multiplesklerose.de/wp-content/uploads/2021/09/KKN_2107_LO_medikamentenhandbuch_24092021.pdf.
47. Montalban X, Gold R, Thompson AJ et al.ECTRIMS/EAN guideline on the pharmacological treatment of people with multiple sclerosis. *Eur J Neurol* 2018; 25(2): 215-237. <https://dx.doi.org/10.1111/ene.13536>.
48. Freedman MS, Selchen D, Prat A et al. Managing Multiple Sclerosis: Treatment Initiation, Modification, and Sequencing. *Can J Neurol Sci* 2018; 45(5): 489-503. <https://dx.doi.org/10.1017/cjn.2018.17>.
49. Noseworthy JH, Lucchinetti C, Rodriguez M et al. Multiple sclerosis. *N Engl J Med* 2000; 343(13): 938-952. <https://dx.doi.org/10.1056/NEJM200009283431307>.
50. Marziniak M, Ghorab K, Kozubski W et al. Variations in multiple sclerosis practice within Europe - Is it time for a new treatment guideline? *Mult Scler Relat Disord* 2016; 8: 35-44. <https://dx.doi.org/10.1016/j.msard.2016.04.004>.

51. Rotstein D, Montalban X. Reaching an evidence-based prognosis for personalized treatment of multiple sclerosis. *Nat Rev Neurol* 2019; 15(5): 287-300. <https://dx.doi.org/10.1038/s41582-019-0170-8>.
52. Carlin CS, Higuera L, Anderson S. Improving Patient-Centered Care by Assessing Patient Preferences for Multiple Sclerosis Disease-Modifying Agents: A Stated-Choice Experiment. *Perm J* 2017; 21: 16-102. <https://dx.doi.org/10.7812/TPP/16-102>.
53. Garcia-Dominguez JM, Munoz D, Comellas M et al. Patient preferences for treatment of multiple sclerosis with disease-modifying therapies: a discrete choice experiment. *Patient Prefer Adherence* 2016; 10: 1945-1956. <https://dx.doi.org/10.2147/PPA.S114619>.
54. Lee Mortensen G, Rasmussen PV. The impact of quality of life on treatment preferences in multiple sclerosis patients. *Patient Prefer Adherence* 2017; 11: 1789-1796. <https://dx.doi.org/10.2147/PPA.S142373>.
55. Poulos C, Wakeford C, Kinter E et al. Patient and physician preferences for multiple sclerosis treatments in Germany: A discrete-choice experiment study. *Mult Scler J Exp Transl Clin* 2020; 6(1): 2055217320910778. <https://dx.doi.org/10.1177/2055217320910778>.
56. Wilson LS, Loucks A, Gipson G et al. Patient preferences for attributes of multiple sclerosis disease-modifying therapies: development and results of a ratings-based conjoint analysis. *Int J MS Care* 2015; 17(2): 74-82. <https://dx.doi.org/10.7224/1537-2073.2013-053>.
57. Wingerchuk DM, Carter JL. Multiple sclerosis: current and emerging disease-modifying therapies and treatment strategies. *Mayo Clin Proc* 2014; 89(2): 225-240. <https://dx.doi.org/10.1016/j.mayocp.2013.11.002>.
58. Gehr S, Kaiser T, Kreutz R et al. Suggestions for improving the design of clinical trials in multiple sclerosis-results of a systematic analysis of completed phase III trials. *EPMA J* 2019; 10(4): 425-436. <https://dx.doi.org/10.1007/s13167-019-00192-z>.
59. Kalincik T, Cutter G, Spelman T et al. Defining reliable disability outcomes in multiple sclerosis. *Brain* 2015; 138(Pt 11): 3287-3298. <https://dx.doi.org/10.1093/brain/awv258>.
60. Huisman E, Papadimitropoulou K, Jarrett J et al. Systematic literature review and network meta-analysis in highly active relapsing-remitting multiple sclerosis and rapidly evolving severe multiple sclerosis. *BMJ Open* 2017; 7(3): e013430. <https://dx.doi.org/10.1136/bmjopen-2016-013430>.
61. Tramacere I, Del Giovane C, Salanti G et al. Immunomodulators and immunosuppressants for relapsing-remitting multiple sclerosis: a network meta-analysis. *Cochrane Database Syst Rev* 2015; (9): CD011381. <https://dx.doi.org/10.1002/14651858.CD011381.pub2>.
62. Fogarty E, Schmitz S, Tubridy N et al. Comparative efficacy of disease-modifying therapies for patients with relapsing remitting multiple sclerosis: systematic review and network meta-analysis. *Mult Scler Relat Disord* 2016; 9: 23-30. <https://dx.doi.org/10.1016/j.msard.2016.06.001>.

63. Li H, Hu F, Zhang Y et al. Comparative efficacy and acceptability of disease-modifying therapies in patients with relapsing-remitting multiple sclerosis: a systematic review and network meta-analysis. *J Neurol* 2020; 267(12): 3489-3498. <https://dx.doi.org/10.1007/s00415-019-09395-w>.
64. Gemeinsamer Bundesausschuss. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V – Ocrelizumab [online]. 2018 [Zugriff: 30.10.2020]. URL: https://www.g-ba.de/downloads/40-268-5167/2018-08-02_AM-RL-XII_Ocrelizumab_D-332_TrG.pdf.
65. Gemeinsamer Bundesausschuss. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V – Fingolimod (neues Anwendungsgebiet) [online]. 2016 [Zugriff: 30.10.2020]. URL: https://www.g-ba.de/downloads/40-268-3789/2016-05-19_AM-RL-XII_Fingolimod_nAWG_D-198_TrG.pdf.
66. ICH Expert Working Group. ICH harmonised tripartite guideline: structure and content of clinical study reports; E3 [online]. 1995 [Zugriff: 18.10.2019]. URL: https://database.ich.org/sites/default/files/E3_Guideline.pdf.
67. Moher D, Hopewell S, Schulz KF et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: c869. <https://dx.doi.org/10.1136/bmj.c869>.
68. Bucher HC, Guyatt GH, Griffith LE et al. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997; 50(6): 683-691. [https://dx.doi.org/10.1016/s0895-4356\(97\)00049-8](https://dx.doi.org/10.1016/s0895-4356(97)00049-8).
69. Higgins JPT, Thompson SG, Deeks JJ et al. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327(7414): 557-560. <https://dx.doi.org/10.1136/bmj.327.7414.557>.
70. Dias S, Welton NJ, Caldwell DM et al. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010; 29(7-8): 932-944. <https://dx.doi.org/10.1002/sim.3767>.
71. Rücker G. Network meta-analysis, electrical networks and graph theory. *Res Syn Meth* 2012; 3(4): 312-324. <https://dx.doi.org/10.1002/jrsm.1058>.
72. Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley; 2008. S. 243-296.