

Systematische Confounderidentifikation in der Indikation schubförmig remittierende multiple Sklerose (RRMS)

A horizontal bar composed of 18 rectangular segments of varying shades of blue and grey. The word 'ARBEITSPAPIER' is centered in white text on a dark blue segment.

ARBEITSPAPIER

Projekt: GA23-02

Version: 1.0

Stand: 30.04.2025

IQWiG-Berichte – Nr. 1994

DOI: 10.60584/GA23-02

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Systematische Confounderidentifikation in der Indikation schubförmig remittierende multiple Sklerose (RRMS)

Auftraggeber

Bearbeitung im Rahmen des Generalauftrags

Interne Projektnummer

GA23-02

DOI-URL

<https://doi.org/10.60584/GA23-02>

Anschrift des Herausgebers

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Siegburger Str. 237
50679 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

ISSN: 1864-2500

Zitiervorschlag

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Systematische Confounderidentifikation in der Indikation schubförmig remittierende multiple Sklerose (RRMS); Arbeitspapier [online]. 2025 [Zugriff: TT.MM.JJJJ]. URL: <https://doi.org/10.60584/GA23-02>.

Schlagwörter

Multiple Sklerose – Schubförmige Remittierende, Propensity Score, Nutzenbewertung, Durchführbarkeitsstudien

Keywords

Multiple Sclerosis – Relapsing-Remitting, Propensity Score, Benefit Assessment, Feasibility Studies

Dieses Arbeitspapier wurde unter Beteiligung externer Sachverständiger erstellt.

Für die Inhalte des Berichts ist allein das IQWiG verantwortlich.

Externe Sachverständige, die wissenschaftliche Forschungsaufträge für das Institut bearbeiten, haben gemäß § 139b Abs. 3 Satz 2 Sozialgesetzbuch – Fünftes Buch – Gesetzliche Krankenversicherung „alle Beziehungen zu Interessenverbänden, Auftragsinstituten, insbesondere der pharmazeutischen Industrie und der Medizinprodukteindustrie, einschließlich Art und Höhe von Zuwendungen“ offenzulegen. Das Institut hat von allen Sachverständigen ein ausgefülltes „Formblatt zur Offenlegung von Beziehungen“ erhalten. Die Angaben wurden durch das speziell für die Beurteilung der Interessenkonflikte eingerichtete Gremium des Instituts bewertet. Die Selbstangaben der externen Sachverständigen zur Offenlegung von Beziehungen sind in Anhang E dargestellt. Es wurden keine Interessenkonflikte festgestellt, die die fachliche Unabhängigkeit im Hinblick auf eine Bearbeitung des vorliegenden Auftrags gefährden.

Externe Sachverständige

- Lars G. Hemkens, Research Center for Clinical Neuroimmunology and Neuroscience Basel (RC2NB), University of Basel and University Hospital Basel
- Julian Hirt, RC2NB, University of Basel and University Hospital Basel
- Perrine Janiaud, RC2NB, University of Basel and University Hospital Basel
- Tim Wölfle, RC2NB, University of Basel and University Hospital Basel

Das IQWiG dankt den externen Beteiligten für ihre Mitarbeit am Projekt.

Mitarbeiterinnen und Mitarbeiter des IQWiG

- Jana Göbel
- Lars Beckmann
- Simone Johner
- Claudia Kapp
- Anne-Kathrin Petri
- Daniela Preukschat
- Barbara Spix
- Volker Vervölgyi
- Siw Waffenschmidt

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	vi
Abbildungsverzeichnis	vii
Abkürzungsverzeichnis.....	viii
Glossar	ix
Kurzfassung	xi
1 Hintergrund.....	1
2 Fragestellung.....	3
3 Projektverlauf.....	4
4 Methoden	5
4.1 Confounderidentifikation gemäß Pufulete 2022 für die Indikation RRMS	5
4.1.1 Confounderidentifikation durch externe Sachverständige.....	5
4.1.2 Einordnung und Kommentierung des Berichts der externen Sachverständigen	5
4.1.3 Überprüfung der Machbarkeit einer Confounderidentifikation gemäß Pufulete 2022	6
4.2 Eigene Untersuchungen.....	6
4.2.1 Zusammenführung der identifizierten potenziellen Confounder	6
4.2.2 Fokussierte Informationsbeschaffung zu Publikationen zur systematischen Vorgehensweise der Confounderidentifikation	6
4.3 Ableitung von Anforderungen an eine systematische Confounderidentifikation (Handlungsempfehlungen)	7
5 Ergebnisse	8
5.1 Confounderidentifikation gemäß Pufulete 2022 für die Indikation RRMS	8
5.1.1 Ergebnis der Confounderidentifikation nach Pufulete 2022	8
5.1.2 Diskussion des Vorgehens der externen Sachverständigen.....	9
5.1.3 Machbarkeit einer Confounderidentifikation gemäß Pufulete 2022.....	14
5.2 Eigene Untersuchungen.....	15
5.2.1 Potenzielle Confounder für die Indikation RRMS.....	15
5.2.2 Weitere Vorschläge zur systematischen Confounderidentifikation in der Literatur	24
5.3 Vorschlag für die Anforderung an eine systematische Confounderidentifikation .	26
5.3.1 Schritt 1: Identifikation.....	28
5.3.2 Schritt 2: Zusammenführung.....	31

5.3.3 Schritt 3: Relevanzbewertung	35
6 Diskussion	36
7 Fazit	40
8 Literatur	42
Anhang A Informationsbeschaffung.....	47
A.1 Liste der ausgeschlossenen Publikationen aus der fokussierten Informationsbeschaffung zur systematischen Vorgehensweise der Confounderidentifikation	47
A.2 Suchstrategien.....	48
A.2.1 Bibliografische Datenbanken	48
Anhang B Zusammenführung der identifizierten potenziellen Confounder.....	49
Anhang C Handlungsempfehlungen – Checkliste	57
Anhang D Externe Expertise zum IQWiG-Arbeitspapier	59
Anhang E Offenlegung von Beziehungen der externen Sachverständigen	128

Tabellenverzeichnis

	Seite
Tabelle 1: Identifiziertes Set potenzieller Confounder inklusive vorgeschlagener Operationalisierungen in der Indikation RRMS	20
Tabelle 2: Zuordnung der identifizierten potenziellen Confounder-Variablen zu den jeweiligen potenziellen Confoundern	49
Tabelle 3: Identifizierte Variablen sortiert nach dem jeweiligen Ausschlussgrund.....	55

Abbildungsverzeichnis

	Seite
Abbildung 1: Grafische Darstellung (Directed Acyclic Graph [DAG]) der Abhängigkeiten von Confounder, Mediator und Collider mit der Intervention und dem zu untersuchenden Endpunkt	x
Abbildung 2: Ergebnis der Zusammenführung der potenziellen Confounder-Variablen	16
Abbildung 3: Ergebnis der fokussierten Informationsbeschaffung und der Studienselektion nach Publikationen zur systematischen Vorgehensweise der Confounderidentifikation	25
Abbildung 4: Vorgeschlagenes 3-schrittiges Verfahren der systematischen Confounderidentifikation	27

Abkürzungsverzeichnis

Abkürzung	Bedeutung
DAG	Directed acyclic Graph (gerichteter azyklischer Graph)
EDSS	Expanded Disability Status Scale
G-BA	Gemeinsamer Bundesausschuss
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
MS	multiple Sklerose
PICO	Population, Intervention, Comparison, Outcome
pU	pharmazeutischer Unternehmer
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
RRMS	Relapsing-remitting multiple Sclerosis (schubförmig remittierende multiple Sklerose)
SUE	schwerwiegendes unerwünschtes Ereignis
UE	unerwünschtes Ereignis

Glossar

Für weitergehende Literatur zu den im Folgenden beschriebenen Begriffsdefinitionen eignen sich z. B. die Publikationen Feeney 2024 [1], Williams 2018 [2], Catalogue of bias collaboration [3], Sendor 2022 [4].

Confounder

Ein Confounder ist eine Variable, die sowohl die Therapieentscheidung als auch den Endpunkt beeinflusst (siehe Abbildung 1) und somit einen Behandlungseffekt verfälschen kann. In nicht randomisierten vergleichenden Studien soll durch die Adjustierung für Confoundern die fehlende Strukturgleichheit zwischen den Behandlungsgruppen ausgeglichen werden.

Nicht messbare Confounder

Unter nicht messbaren Confoundern werden in diesem Bericht solche Confounder zusammengefasst, die zwar bekannt sind, aber nicht operationalisiert und entsprechend nicht erhoben werden können, wie z. B. der Patientenwunsch oder die Arztentscheidung. Darüber hinaus bleiben selbst bei sorgfältiger und systematischer Confounderidentifikation potenziell gänzlich unbekannte Confounder bestehen.

Diese Confounder sind abzugrenzen von solchen Confoundern, die eindeutig operationalisiert werden können, aber in der zur Verfügung stehenden Datenquelle nicht verfügbar sind (siehe auch residual confounding).

Residuelles Confounding / Unmeasured Confounding

Unter den Begriffen Residual Confounding und Unmeasured Confounding werden in der Literatur 3 Situationen verstanden, in denen die Analysen zu verzerrten Ergebnissen führen können. Erstens, für einen bekannten Confounder liegen keine Werte vor, bzw. der Confounder kann nicht beobachtet und operationalisiert werden. Zweitens, für einen bekannten Confounder liegen zwar beobachtete Werte vor, diese sind aber fehlerhaft beobachtet. In der dritten Situation ist ein Confounder nicht bekannt. Unter residuellem Confounding wird somit die Gesamtheit der nicht gemessenen und somit nicht in der statistischen Analyse berücksichtigten Confounder verstanden.

Confounding by Indication

Confounding by Indication ergibt sich dann, wenn eine Variable, z. B. ein Symptom oder prognostischer Faktor, die die Therapieentscheidungen beeinflusst, gleichzeitig auch auf die Ergebnisse der untersuchten Endpunkte Einfluss hat. Ein klassisches Beispiel hierfür: eine Erkrankung [oder die Schwere der Erkrankung] ist ein Risikofaktor für den zu untersuchenden Endpunkt. Wird die Erkrankung mit der Studienintervention behandelt, dann haben die

behandelten Patientinnen und Patienten allein aufgrund ihrer von vornherein schwereren Erkrankung ein höheres Risiko für ein schlechtes Ergebnis im zu untersuchenden Endpunkt als die Nichtbehandelten.

Gerichteter azyklischer Graph (Directed Acyclic Graph [DAG])

DAGs sind ein grafisches Werkzeug, das eine Möglichkeit bietet, die (angenommenen) kausalen Zusammenhänge und Abhängigkeiten von Variablen visuell darzustellen und besser zu verstehen.

Collider

Ein Collider wird sowohl von der Intervention als auch dem Endpunkt beeinflusst (siehe Abbildung 1). Die Adjustierung für einen Collider ist zu vermeiden, da dies eine Scheinkorrelation zwischen Intervention und Endpunkt erzeugen kann.

Mediator

Ein Mediator liegt auf dem direkten Weg zwischen der Intervention und dem Endpunkt. Die Adjustierung für einen Mediator ist ebenfalls zu vermeiden.

Target Trial Emulation

Unter einer Target Trial Emulation wird die Nachbildung der Planung von vergleichenden Studien mit Randomisierung auf Basis von Beobachtungsdaten verstanden.

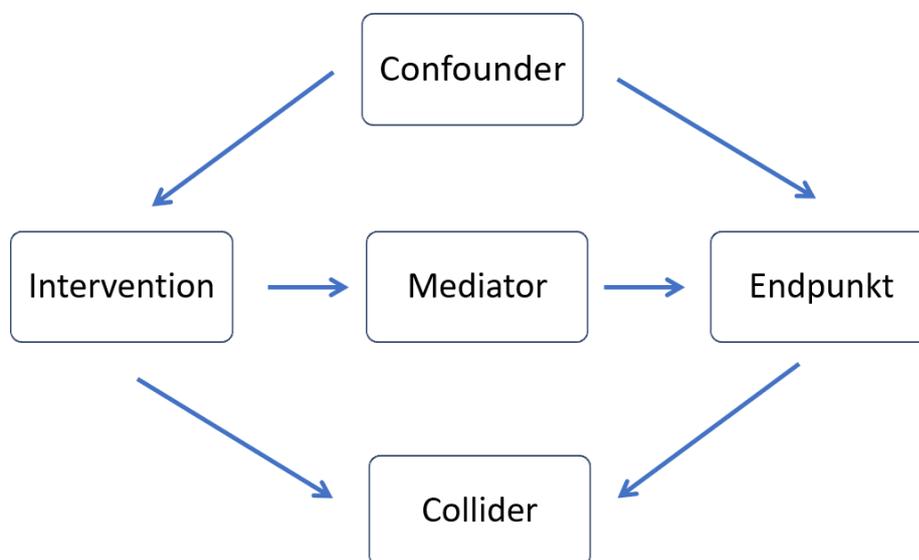


Abbildung 1: Grafische Darstellung (Directed Acyclic Graph [DAG]) der Abhängigkeiten von Confounder, Mediator und Collider mit der Intervention und dem zu untersuchenden Endpunkt

Kurzfassung

Im Rahmen des Generalauftrags wurde das Thema Systematische Confounderidentifikation in der Indikation schubförmig remittierende multiple Sklerose (RRMS) bearbeitet.

Hintergrund

In nicht randomisierten Vergleichen von Interventionen ist für eine adäquate Confounderkontrolle eine systematische Identifikation der relevanten Confounder sowie eine Berücksichtigung dieser Confounder in der Auswertung notwendig. Die Publikation Pufulete 2022 stellt bisher den einzigen veröffentlichten konkreten Vorschlag für eine systematische Confounderidentifikation dar. Nach Pufulete 2022 erfolgt die Confounderidentifikation über 3 methodische Bausteine: systematische Literaturrecherche, Interviews mit Klinikerinnen und Klinikern und eine Umfrage unter Klinikerinnen und Klinikern.

Mit dem vorliegenden Projekt soll eine systematische Confounderidentifikation in der Indikation der RRMS durchgeführt werden, um damit als Modellprojekt zu dienen, das die grundsätzliche Machbarkeit eines Vorgehens zur systematischen Confounderidentifikation gemäß Pufulete 2022 überprüft. Darüber hinaus sollen konkrete Handlungsempfehlungen zur systematischen Identifikation von Confoundern für nicht randomisierte vergleichende Studien und somit auch für die anwendungsbegleitende Datenerhebung abgeleitet werden.

Fragestellung

Ziele der vorliegenden Untersuchung sind die

- systematische Confounderidentifikation gemäß Pufulete 2022 für die Indikation RRMS an einem geeigneten Vergleich zweier medikamentöser Therapien hinsichtlich wesentlicher patientenrelevanter Endpunkte
- Überprüfung der grundsätzlichen Machbarkeit einer Confounderidentifikation gemäß Pufulete 2022
- Ableitung von Handlungsempfehlungen zur systematischen Identifikation von Confoundern für nicht randomisierte vergleichende Studien

Methoden

Im vorliegenden Projekt wurden externe Sachverständige mit einer systematischen Confounderidentifikation für die Indikation RRMS gemäß der Publikation Pufulete 2022 beauftragt. Pufulete 2022 beschreibt eine Methodik für eine solche systematische Confounderidentifikation.

Der Bericht der externen Sachverständigen wurde von der IQWiG-Projektgruppe eingeordnet und kommentiert. Zusätzlich wurde anhand dieses Berichts, dem beanspruchten Zeitrahmen und den eingesetzten Ressourcen die Machbarkeit einer Confounderidentifikation gemäß Pufulete 2022 überprüft.

Basierend auf dem Bericht der externen Sachverständigen wurden weitere eigene Untersuchungen durchgeführt, die über den methodischen Ansatz von Pufulete 2022 hinausgehen. Es wurde geprüft, ob die im Bericht der externen Sachverständigen identifizierten potenziellen Confounder-Variablen weitergehend zusammengeführt werden können.

Darüber hinaus wurde anhand einer fokussierten Informationsbeschaffung geprüft, ob neben dem von Pufulete 2022 vorgeschlagenen systematischen Ansatz zur Confounderidentifikation weitere Publikationen zur systematischen Confounderidentifikation vorliegen.

Abschließend wurden anhand der Publikation von Pufulete 2022, der Erfahrungen aus dem Bericht der externen Sachverständigen zur Identifikation von Confoundern gemäß Pufulete 2022 und anhand der weiterführenden eigenen Untersuchungen konkrete Handlungsempfehlungen zur systematischen Identifikation von Confoundern für nicht randomisierte vergleichende Studien abgeleitet. Hierbei wurde im Speziellen geprüft, inwieweit sich für Fragestellungen mit erwartbar weniger umfassender Studienlage (wie beispielsweise Fragestellungen der anwendungsbegleitenden Datenerhebung) besondere Anforderungen ergeben.

Ergebnisse

Confounderidentifikation gemäß Pufulete 2022 für die Indikation RRMS

Ergebnis der Confounderidentifikation nach Pufulete 2022 und Diskussion des Vorgehens der externen Sachverständigen

Die externen Sachverständigen legen einen ausführlichen Bericht mit einer umfassenden Confounder-Recherche nach Pufulete 2022 vor.

Die externen Sachverständigen haben über den systematischen 3-stufigen Ansatz der Confounderidentifikation 160 potenzielle Confounder identifiziert. 125 potenzielle Confounder wurden dabei über die Literaturrecherche (Suche nach Beobachtungsstudien und RCTs) identifiziert. Über die Interviews (n = 8) mit Klinikerinnen und Klinikern aus Deutschland und aus der Schweiz wurden insgesamt 53 potenzielle Confounder identifiziert, 18 davon wurden auch durch die Literaturrecherche ermittelt, 35 wurden erstmals durch die Klinikerinnen und Kliniker genannt. Im anschließenden quantitativen Survey mit 21 Personen (von denen 7 die Umfrage beendeten) wurden die identifizierten Confounder dahin gehend beurteilt, wie stark diese die Therapieentscheidung der Ärztin oder des Arztes beeinflussen.

Die externen Sachverständigen haben hierfür die 160 identifizierten Confounder in ersten Schritten zu 137 Confoundern zusammengefasst und reduziert (13 über ähnliche Confounder abgedeckt, 2 ohne Studienkontext nicht beurteilbar, 8 Confounder waren zu Beginn der Umfrage noch nicht identifiziert). Die Einstufung erfolgte mittels einer Skala von "1 – keine Berücksichtigung" bis "5 – sehr starke Berücksichtigung. 40 der 137 Confounder wurden im Mittel mit einem Wert > 4 bewertet.

Die externen Sachverständigen haben 148 (93 %) der 160 potenziellen Confounder als „wahrscheinlich ein Prädiktor für die Behandlung“ und von diesen 148 potenziellen Confoundern 128 (80 % der Gesamtzahl) als „wahrscheinlich auch ein Prädiktor für die betrachteten Endpunkte“ eingestuft. 11 (7 %) potenzielle Confounder wurden als unklar eingestuft, was die externen Sachverständigen als „wahrscheinlich weder ein Prädiktor für die Behandlung noch für die Endpunkte“ interpretieren. Für den potenziellen Confounder Type of MS wurde keine Einstufung durch die externen Sachverständigen vorgenommen.

In einem letzten Schritt wurden die 160 potenziellen Confounder von den externen Sachverständigen zu 136 Confoundern geclustert (d. h. Confounder, die auf ähnliche Weise gemessen oder bewertet werden können).

Insgesamt bearbeiteten die externen Sachverständigen die Fragestellung angemessen entsprechend den Vorgaben in Pufulete 2022. Im Folgenden wird beispielhaft auf Abweichungen zu dem Vorgehen gemäß Pufulete 2022 eingegangen.

Abweichend zu Pufulete 2022, die bei der Extraktion der Confounder aus den systematisch identifizierten Literaturquellen ein Sättigungskriterium verwendeten, legten die externen Sachverständigen ein Sättigungskriterium von 15 konsekutiven Studien nur für Studien vor dem Jahr 2020 an. Von den relevanten Studien, die seit 2020 publiziert wurden, wurden hingegen alle extrahiert. Anhand der vorliegenden Informationen im Bericht der externen Sachverständigen kann nicht eingeschätzt werden, inwiefern der Mehraufwand den zusätzlichen Gewinn an Confoundern (57/125 [46 %] wurden ausschließlich in Studien identifiziert, die 2020 oder später veröffentlicht wurden) rechtfertigt. Die externen Sachverständigen gehen davon aus, eine vollständigere Liste identifiziert zu haben, aber beschreiben den gewählten Ansatz selbst als sehr ressourcenintensiv.

Hinsichtlich der qualitativen Interviews gehen Pufulete 2022 in ihrer Publikation davon aus, dass mit 12 Expertinnen und Experten (je 6 in Kardiologie und Herzchirurgie) eine Sättigung erreicht werden kann. Aufgrund von Schwierigkeiten bei der Rekrutierung wurden von den externen Sachverständigen letztlich nur 8 statt der geplanten 10 klinischen Expertinnen und Experten eingebunden. Die externen Sachverständigen geben an, dass die Beiträge der 8 Expertinnen und Experten gesättigt waren und die bei der Therapieentscheidung berücksichtigten Faktoren ausreichend widerspiegeln. Die Einbindung von 8 klinischen

Expertinnen und Experten für die qualitativen Interviews wird insgesamt als angemessen beurteilt. Ob ggf. auch eine geringere Zahl an Interviews bereits eine Sättigung erzielt hätte, kann an dieser Stelle nicht beurteilt werden.

In der quantitativen Umfrage sollten die Teilnehmenden der Umfrage nicht wie bei Pufulete 2022 die 5 wichtigsten Confounder benennen, sondern auf einer 5-Punkte-Skala mit einem Score von 1 – keine Berücksichtigung bis 5 – starke Berücksichtigung bewerten, wie stark jeder einzelne potenzielle Confounder die Therapieentscheidung beeinflusst. Dies stellt eine alternative Herangehensweise dar, die Relevanz der Confounder einzuschätzen. Auch wenn von den beiden möglichen Bewertungsmethoden das Scoring einen höheren Aufwand bedeutet, ist die Relevanzbewertung eines jeden einzelnen Confounders im Vergleich zu dem von Pufulete 2022 gewählten Vorgehen der Benennung der 5 wichtigsten Confounder in der Gesamtschau inhaltlich zu befürworten. Mithilfe der Scores können alle potenziellen Confounder hinsichtlich ihrer Bedeutung sortiert werden. Diese Rangfolge kann nachgelagert ggf. als Entscheidungshilfe genutzt werden (z. B. in Fällen in denen das statistische Modell nicht konvergiert kann auf Basis der Scores begründet werden, welcher potenzielle Confounder bei der Adjustierung im Modell als erstes verzichtbar erscheint). Um den Aufwand für diese Confounder-Bewertung zu reduzieren, wird ein möglicher Ansatz der intensiveren Kondensation der Liste der zu bewertenden potenziellen Confounder vorgeschlagen (siehe unten).

In einem letzten Schritt haben sowohl Pufulete 2022 als auch die externen Sachverständigen beurteilt, ob der potenzielle Confounder sowohl die Therapieentscheidung als auch das Behandlungsergebnis beeinflusst. Das Ziel einer solchen Kategorisierung ist, „wahre Confounder“ zu identifizieren. Da die Einschätzung, in welchem Zusammenhang die Variablen stehen, auf starken Annahmen beruht und entsprechend mit Unsicherheiten behaftet ist, ist eine solche Identifikation der Confounder in der Regel nicht möglich. Entsprechend sollten Confounder-Variablen, deren Einfluss auf Therapieentscheidung und Behandlungsergebnis unsicher ist, nicht ausgeschlossen werden. Lediglich eine literaturgestützte Begründung kann den Ausschluss eines potenziellen Confounders gegebenenfalls rechtfertigen.

Machbarkeit einer Confounderidentifikation gemäß Pufulete 2022

Die Prüfung des Berichts der externen Sachverständigen ergab, dass im Wesentlichen adäquat gemäß den Vorgaben in Pufulete 2022 vorgegangen wurde. Grundsätzlich wird die Durchführung der Confounderidentifikation nach dem Ansatz von Pufulete 2022 somit als machbar bewertet.

Die externen Sachverständigen haben für die Identifikation der 160 Confounder 62 Personentage (à 8 Arbeitsstunden) benötigt. Die systematische Literaturrecherche stellt mit 33 geschätzten Personentagen den zeitaufwendigsten Baustein der Confounderidentifikation

dar. Hierbei ist zu berücksichtigen, dass die Indikation RRMS unter anderem aus dem Grund ausgewählt wurde, dass insgesamt von einer ausreichenden Studienlage ausgegangen wird. In Indikationen mit erwartbar weniger umfassender Studienlage (und entsprechend einer deutlich geringeren Anzahl an zu screenenden und zu extrahierenden Treffern der Literaturrecherche) kann der Zeitaufwand für diesen Baustein deutlich geringer ausfallen. Darüber hinaus ist der von den externen Sachverständigen gewählte Ansatz, für Publikationen nach 2020 kein Sättigungskriterium für die Confounderidentifikation zu wählen, sehr sensitiv. In diesem Schritt besteht somit ggf. Einsparpotenzial durch die Verwendung eines Sättigungskriteriums für alle Publikationen.

Eigene Untersuchungen

Potenzielle Confounder für die Indikation RRMS

Die externen Sachverständigen identifizierten 160 potenzielle Confounder-Variablen und clusterten Confounder, die auf ähnliche Weise gemessen oder bewertet werden können zu 136 potenziellen Confoundern. Eine Zusammenführung der über die Literaturrecherche und die Interviews mit klinischen Expertinnen und Experten identifizierten Confounder war nicht Bestandteil der Ausarbeitung von Pufulete 2022, wird von den Autorinnen und Autoren aber als logischen weiteren Schritt diskutiert. Insgesamt gibt es für diesen Schritt aber bislang keine Handlungsempfehlungen.

Es wurde überprüft, inwieweit eine weitergehende intensivere Zusammenführung der identifizierten potenziellen Confounder-Variablen vorgenommen werden kann, die höchstens mit einem vertretbaren Informationsverlust hinsichtlich der identifizierten potenziellen Confounder einhergeht. Die vorliegende systematische Confounderidentifikation soll zur Durchführung eines nicht randomisierten Vergleichs der Interventionen Dimethylfumarat vs. Glatirameracetat bei Patientinnen und Patienten ab 18 Jahren in der Indikation RRMS genutzt werden können. Im Folgenden wird angenommen, dass die RCT CONFIRM als Modellstudie mittels Target Trial Emulation nachgebildet werden soll. Somit konnten die konkreten Ein- und Ausschlusskriterien der RCT CONFIRM genutzt werden, um Variablen, die für die vorliegende Fragestellung keine Relevanz haben, auszuschließen. Insgesamt wurden 28 von 160 Variablen basierend auf der vorliegenden Fragestellung ausgeschlossen.

Die 132 verbleibenden potenziellen Confounder-Variablen wurden mit dem Ziel, Duplikate und Überlappungen zu reduzieren, zusammengeführt. Beispielsweise wurden dem Confounder gesundheitsbezogene Lebensqualität (Quality of Life) die potenziellen Confounder-Variablen EQ-5D, SF-36 Scores, Multiple Sclerosis Impact Scale, Multiple Sclerosis International Questionnaire of Quality of Life und Sexual Satisfaction Scale Score zugeordnet. Insgesamt konnten die 132 potenziellen Confounder-Variablen zu 28 potenziellen Confoundern verdichtet werden.

Von diesen 28 potenziellen Confoundern konnten 26 mithilfe der Literatur und 18 über die Interviews mit klinischen Expertinnen und Experten identifiziert werden. Die potenziellen Confounder Patientenwunsch und Art der beruflichen Tätigkeit wurden nicht anhand der Literatur, sondern nur aufgrund der Interviews identifiziert.

Operationalisierung der potenziellen Confounder

Zuletzt wurden geeignete Operationalisierungen der potenziellen Confounder vorgeschlagen. Falls möglich, wurde eine stetige Operationalisierung der Confounder bevorzugt, da diese den höchsten Informationsgehalt bietet. Sind stetige Operationalisierungen nicht möglich, wird eine kategorielle Operationalisierung gewählt. Nur für den Fall, dass statistische Modelle nicht konvergieren, sollten mögliche Zusammenfassungen von Ausprägungen (z. B. Dichotomisierungen) in Betracht gezogen werden. Diese Zusammenfassungen sind zu präzisieren.

Für die beiden Confounder Patientenwunsch sowie Verträglichkeit der Vortherapie und Erwartungen an die Behandlung war aufgrund der Komplexität und Vielschichtigkeit keine adäquate Operationalisierung möglich. Daher können die beiden als relevant eingestuften potenziellen Confounder nicht in die Auswertung einbezogen werden. Dies muss anschließend bei der Ergebnisinterpretation berücksichtigt werden.

Weitere Vorschläge zur systematischen Confounderidentifikation in der Literatur

Neben der Publikation Pufulete 2022 selbst konnten keine weiteren Publikationen zu systematischen Ansätzen der Confounderidentifikation identifiziert werden.

Vorschlag für die Anforderung an eine systematische Confounderidentifikation

Zur Evaluation von Interventionseffekten auf Basis eines nicht randomisierten Vergleichs, sollte im Rahmen der Studienplanung a priori ein Set potenzieller Confoundern identifiziert werden. Hierfür wurden basierend auf den Ergebnissen einer Confounderidentifikation durch externe Sachverständige und anhand der weitergeführten eigenen Untersuchungen konkrete Handlungsempfehlungen zur systematischen Identifikation von Confoundern abgeleitet. Diese stellen einen Vorschlag für die konkreten Anforderungen an die systematische Confounderidentifikation dar.

Auch wenn sich der Bericht der externen Sachverständigen (ebenso wie auch die Arbeit von Pufulete 2022) auf eine Indikation mit umfangreicher Studienlage bezieht, wurde bei der Ableitung der Handlungsempfehlungen auch die Situation der selten(er)en Erkrankungen berücksichtigt, da dies die regelhafte Situation z. B. im Rahmen der anwendungsbegleitenden Datenerhebung darstellt. Sollten sich hierfür besondere Anforderungen ergeben, wird dies im Folgenden explizit beschrieben.

Übergeordnet gliedert sich der erarbeitete Vorschlag zur systematischen Confounderidentifikation in 3 Schritte.

Schritt 1: Identifikation mittels systematischer Literaturrecherche und Interviews mit klinischen Expertinnen und Experten

Angelehnt an Pufulete 2022 setzt sich der 1. Schritt der Identifikation aus 2 Bausteinen zusammen: einer systematischen Literaturrecherche und Interviews mit klinischen Expertinnen und Experten. Bei der systematischen Informationsbeschaffung sollte auf Primärpublikationen zu RCTs und Kohortenstudien zurückgegriffen werden. Nachfolgend können auch Leitlinien oder 1-armige Studien berücksichtigt werden. Die Ein- und Ausschlusskriterien für den Studienpool zur Confounderextraktion sollten sich nah an der zu untersuchenden Fragestellung orientieren. Die Extraktion sollte bei vergleichenden Beobachtungsstudien alle Adjustierungsvariablen – unabhängig von der statistischen Signifikanz – umfassen. Bei randomisierten oder nicht vergleichenden Studien sollten alle Patientencharakteristika extrahiert werden. Bei der Extraktion aus Primärpublikationen wird die Anwendung eines geeigneten Sättigungskriteriums empfohlen. Als 2. Baustein sollten Interviews mit klinischen Expertinnen und Experten durchgeführt werden. Die klinischen Expertinnen und Experten sollten namentlich benannt und potenzielle Interessenkonflikte offengelegt werden.

Schritt 2: Zusammenführung

Nach der erfolgten Extraktion aller potenziellen Confounder-Variablen wird im 2. Schritt eine Zusammenführung inhaltlich überlappender Confounder-Variablen vorgenommen. Für jede identifizierte Confounder-Variable sollte geprüft werden, ob sie innerhalb der zu untersuchenden Fragestellung liegt. Der Ausschluss eines potenziellen Confounders aus inhaltlichen Überlegungen muss literaturgestützt begründet werden. Eine transparente und nachvollziehbare Dokumentation sollte gewährleistet sein. Für alle identifizierten potenziellen Confounder ist eine Operationalisierung festzulegen. Nicht messbare bzw. nicht operationalisierbare potenzielle Confounder können zwar nicht in die Auswertung einbezogen werden, sollten jedoch benannt und bei der Ergebnisinterpretation berücksichtigt werden.

Schritt 3: Relevanzbewertung der potenziellen Confounder durch klinische Expertinnen und Experten

Anschließend erfolgt im 3. Schritt eine quantitative Bewertung der bereits zusammengeführten, potenziellen Confounder hinsichtlich ihrer Relevanz durch die klinischen Expertinnen und Experten. Eine Bewertung mittels Punktevergabe wird empfohlen.

Dokumentation der Confounderidentifikation

Um die Confounderidentifikation nachvollziehen zu können, ist eine detaillierte und transparente Dokumentation der Vorgehensweise und der Ergebnisse erforderlich. Diese

Dokumentation sollte als Teil des Studienprotokolls der geplanten nicht randomisierten vergleichenden Studie veröffentlicht werden.

Fazit

Die externen Sachverständigen haben unter Anwendung des von Pufulete 2022 vorgeschlagenen Vorgehens insgesamt 160 potenzielle Confounder-Variablen identifiziert. Diese wurden in weiterführenden eigenen Untersuchungen zu 28 potenziellen Confoundern zusammengeführt. Damit konnte demonstriert werden, dass die Confounderidentifikation gemäß Pufulete 2022 grundsätzlich machbar ist, z. B. im Rahmen einer anwendungsbegleitenden Datenerhebung.

Aus den in den Untersuchungen gewonnenen Erfahrungen wurden die folgenden Handlungsempfehlungen für eine systematische Confounderidentifikation in 3 Schritten abgeleitet:

1) Confounderidentifikation

Potenzielle Confounder-Variablen sollen aus einer systematischen Literaturrecherche und aus Interviews mit klinischen Expertinnen und Experten identifiziert werden.

2) Confounder-Zusammenführung

Confounder-Variablen, welche ein inhaltlich gleiches Konstrukt messen oder sich inhaltlich überlappen, sollen zusammengeführt werden. Für jeden dieser potenziellen Confounder ist eine eindeutige Operationalisierung festzulegen.

3) Confounder-Bewertung

Jeder potenzielle Confounder sollte über eine Umfrage mit klinischen Expertinnen und Experten in seiner Bedeutung bewertet werden. Das Ergebnis kann für eine inhaltlich gestützte Variablenselektion verwendet werden.

Insbesondere der 2. Schritt stellt dabei eine Weiterentwicklung der von Pufulete 2022 vorgeschlagenen Methodik dar, mit dem Ziel, den Aufwand ohne relevanten Informationsverlust zu reduzieren.

Das Verfahren der systematischen Confounderidentifikation ist trotz der beschriebenen Möglichkeiten zur Aufwandseinsparung ressourcenintensiv. In einem nächsten Schritt soll geprüft werden, welche weiteren möglichen Optionen zur Aufwandsreduktion bestehen, sowohl für die Confounderidentifikation als auch die zugehörige Erhebung der potenziellen Confounder. Die möglichen Optionen zur Aufwandsreduktion müssen dem damit einhergehenden Informationsverlust sowie dem potenziellen Verlust der Ergebnissicherheit gegenübergestellt werden.

Um in einer nicht randomisierten vergleichenden Studie die Verzerrung durch Confounding möglichst zu minimieren, ist eine systematische Confounderidentifikation erforderlich. Dennoch verbleibt in jedem Fall eine Unsicherheit, ob wirklich alle relevanten Confounder identifiziert wurden und dementsprechend in der Analyse berücksichtigt werden können. Nicht bekannte Confounder ebenso wie nicht messbare Confounder bleiben ein ungelöstes Problem, sodass in nicht randomisierten vergleichenden Studien nur bei hinreichend großen Effekten bspw. ein Zusatznutzen einer Therapie ableitbar ist. Eine sinnvolle Alternative ist einzig die Durchführung einer randomisierten vergleichenden Studie, wie z. B. einer registerbasierten RCT, für die keine Confounderidentifikation und -kontrolle erforderlich ist.

1 Hintergrund

Allgemeine Bedeutung der systematischen Confounderidentifikation

Zum Zeitpunkt der Zulassung neuer Medikamente bestehen häufig noch Evidenzlücken, z. B. zum Vergleich verschiedener Therapieoptionen untereinander. Diese können z. B. durch registerbasierte randomisierte kontrollierte Studie (RCT) geschlossen werden. Alternativ werden Studien ohne Randomisierung unter Verwendung einer Target Trial Emulation [5-9] einschließlich sorgfältiger Korrektur für Confounder (z. B. mittels Propensity Score – Verfahren) diskutiert. In nicht randomisierten Vergleichen von Interventionen ist für eine adäquate Confounderkontrolle eine systematische Identifikation der relevanten Confounder sowie eine Berücksichtigung dieser Confounder in der Auswertung notwendig. Da anwendungsbegleitende Datenerhebungen gemäß § 35a Abs.3b SGB V in der Regel vergleichende Studien ohne Randomisierung darstellen, spielt auch bei diesen eine adäquate Adjustierung für Confounder eine zentrale Rolle. Hier ist es erforderlich, vorab alle relevanten Confounder (inklusive wichtiger Wechselwirkungen) mithilfe eines systematischen Ansatzes zu identifizieren, vollständig zu erheben und bei der Datenanalyse in angemessener Form im Modell zu berücksichtigen [10]. Bei den anwendungsbegleitenden Datenerhebungen ergibt sich zudem die besondere Situation, dass sie in der Regel in Anwendungsgebieten mit kleinen Patientenzahlen durchgeführt wird, in denen nur wenig Evidenz vorliegt.

Vorgehen gemäß Pufulete 2022

In der Publikation von Pufulete 2022 [11] ist beschrieben, wie eine Confounderidentifikation über 3 methodische Bausteine (systematische Literaturrecherche, Interviews mit Klinikerinnen und Klinikern, Umfrage unter Klinikerinnen und Klinikern) erfolgen kann. Dazu wurden erstens Daten zu potenziellen Confoundern und Ko-Interventionen unter zur Hilfenahme eines Sättigungskriteriums jeweils unabhängig von 2 Wissenschaftlerinnen und Wissenschaftlern extrahiert, zweitens durch Interviews mit Klinikerinnen und Klinikern weitere potenzielle Confounder identifiziert und darauf in verschiedene Kategorien eingeteilt und drittens anhand der Umfrage jeweils die 5 wichtigsten potenziellen Confounder ermittelt. Zuletzt wurden die identifizierten Confounder entsprechend ihres vermuteten Einflusses auf die Therapieentscheidung und / oder das Behandlungsergebnis klassifiziert und die vermutete Verzerrungsrichtung beschrieben (siehe dazu Tabelle 1 Pufulete 2022). Auch wenn in der Literatur eine systematische Herangehensweise bei der Identifikation von Confoundern gefordert wird [12-14], stellt die Publikation Pufulete 2022 bisher den einzigen veröffentlichten konkreten Vorschlag für eine systematische Confounderidentifikation im Bereich klinischer und epidemiologischer Studien dar. Entsprechend wird sie in diesem Zusammenhang vom IQWiG sowohl in den Allgemeinen Methoden [15] als auch in den Projekten im Rahmen der anwendungsbegleitenden Datenerhebung [16,17] beispielhaft erwähnt.

Rationale für das Projekt

In der Studienplanung für die anwendungsbegleitende Datenerhebung wurde bereits von verschiedenen pharmazeutischen Unternehmen (pU) eine systematische Confounderidentifikation vorgenommen (z. B. [18,19]). Keine der vorgelegten Untersuchungen hat die Anforderungen an die systematische Confounderidentifikation vollumfänglich erfüllt (vgl. z. B. [20-23]). In diesem Zusammenhang wird häufig die Kritik geäußert, dass das von Pufulete 2022 vorgeschlagene Vorgehen zur systematischen Confounderidentifikation mit einem sehr hohen Ressourcenaufwand verbunden sei.

Mit dem vorliegenden Projekt soll eine systematische Confounderidentifikation in der Indikation der schubförmig remittierenden multiplen Sklerose (RRMS) durchgeführt werden, um damit als Modellprojekt zu dienen, das die grundsätzliche Machbarkeit eines Vorgehens zur systematischen Confounderidentifikation gemäß Pufulete 2022 überprüft. Die Indikation RRMS stellt eine Indikation mit einem erwartbar hohen Umfang an verfügbarer Evidenz dar, was der von Pufulete 2022 untersuchten Indikation der koronaren Herzkrankheit entspricht.

Darüber hinaus sollen konkrete Handlungsempfehlungen zur systematischen Identifikation von Confoundern für nicht randomisierte vergleichende Studien und somit auch für die anwendungsbegleitende Datenerhebung abgeleitet werden. Die Besonderheiten von Indikationen mit erwartbar weniger umfassender Studienlage werden hierbei explizit adressiert.

Zudem sollen die im vorliegenden Arbeitspapier identifizierten Confounder in der Indikation RRMS im Anschluss, z. B. im Rahmen eines Folgeprojekts, für die Ableitung von Therapieeffekten auf Basis eines nicht randomisierten Vergleichs genutzt werden können. Dies soll im Rahmen einer Nachbildung einer konkreten RCT durchgeführt werden, wobei die Studie CONFIRM, die bei erwachsenen Patientinnen und Patienten mit RRMS den Vergleich von Dimethylfumarat und Glatirameracetat untersucht [24], als Modellstudie für eine Target Trial Emulation dienen soll.

2 Fragestellung

Ziele der vorliegenden Untersuchung sind die

- systematische Confounderidentifikation gemäß Pufulete 2022 für die Indikation schubförmig remittierende multiple Sklerose (RRMS) an einem geeigneten Vergleich zweier medikamentöser Therapien hinsichtlich wesentlicher patientenrelevanter Endpunkte
- Überprüfung der grundsätzlichen Machbarkeit einer Confounderidentifikation gemäß Pufulete 2022
- Ableitung von Handlungsempfehlungen zur systematischen Identifikation von Confoundern für nicht randomisierte vergleichende Studien

3 Projektverlauf

Im Rahmen des Generalauftrags wurde das Thema Systematische Confounderidentifikation in der Indikation schubförmig remittierende multiple Sklerose (RRMS) bearbeitet. Bearbeitungsbeginn war der 05.05.2023.

In die Bearbeitung des Projekts wurden externe Sachverständige eingebunden und mit der Durchführung einer systematischen Confounderidentifikation gemäß Pufulete 2022 für die Indikation RRMS beauftragt. Basierend auf dem Bericht der externen Sachverständigen wurden zum einen die grundsätzliche Machbarkeit des von Pufulete 2022 vorgeschlagenen Vorgehens beurteilt und zum anderen weitere eigene Untersuchungen durchgeführt, die über den methodischen Ansatz von Pufulete 2022 hinausgehen, um schließlich übergeordnete Handlungsempfehlungen zur systematischen Confounderidentifikation zu erarbeiten.

Auf Basis einer internen Projektskizze wurde ein Arbeitspapier erstellt. Dieser Bericht wurde an den G-BA übermittelt und 4 Wochen später auf der Website des IQWiG veröffentlicht.

4 Methoden

4.1 Confounderidentifikation gemäß Pufulete 2022 für die Indikation RRMS

4.1.1 Confounderidentifikation durch externe Sachverständige

Im vorliegenden Projekt wurden externe Sachverständige mit einer systematischen Confounderidentifikation für die Indikation RRMS gemäß der Publikation Pufulete 2022 [11] beauftragt. Die anzuwendende Methodik für die Confounderidentifikation ist in Pufulete 2022 beschrieben.

Der Auftrag wurde zunächst vollständig und unabhängig von der externen Gruppe durchgeführt, die über entsprechende Expertise im Gebiet nicht randomisierte Studien (Non-RCT) / Confounder verfügt. Im vorliegenden Arbeitspapier wird das methodische Vorgehen und das Ergebnis der externen Sachverständigen vom IQWiG eingeordnet. Dazu wird das methodische Vorgehen im Abgleich mit Pufulete 2022 diskutiert.

Ausgewählte Fragestellung

Die identifizierten Confounder sollen für die Nachbildung der RCT CONFIRM [24] anhand eines nicht randomisierten Vergleichs der Interventionen Dimethylfumarat vs. Glatirameracetat bei Patientinnen und Patienten ab 18 Jahren in der Indikation RRMS genutzt werden können. Die RCT CONFIRM wurde primär deswegen ausgewählt, da sie sich gut für die geplante Target Trial Emulation eignet (u. a. wegen des Vergleichs mit einem aktiven Komparator); zudem ist für die zugrundeliegende Fragestellung von einer insgesamt ausreichenden Studienlage für die Confounderidentifikation auszugehen. Die zentralen patientenrelevanten Endpunkte, für die im nicht randomisierten Vergleich Therapieeffekte mit entsprechender Adjustierung für Confounder untersucht werden sollen, waren: die jährliche Schubrate, die bestätigte Behinderungsprogression anhand der Expanded Disability Status Scale (EDSS), gesundheitsbezogene Lebensqualität und unerwünschte Ereignisse (UEs), insbesondere schwerwiegende unerwünschte Ereignisse (SUE). Die Übertragbarkeit der Machbarkeit einer Confounderidentifikation in der vorliegenden Fragestellung auf andere Fragestellungen wird in Abschnitt 5.3 diskutiert.

4.1.2 Einordnung und Kommentierung des Berichts der externen Sachverständigen

Der Bericht der externen Sachverständigen wurde in Abschnitt 5.1.2 des Arbeitspapiers von der IQWiG-Projektgruppe eingeordnet und kommentiert. Der Bericht der externen Gruppe ist als Anhang dem Arbeitspapier beigefügt (siehe Anhang D). Für die interne Prüfung des methodischen Vorgehens des Berichtes war maßgeblich, ob gemäß den Vorgaben in Pufulete 2022 vorgegangen wurde bzw. ob ggf. vorgenommene Anpassungen, z. B. zur Aufwandsreduzierung, als sinnvoll und vertretbar eingeschätzt werden.

Basierend auf den beiden verfügbaren Anwendungsbeispielen (Confounderidentifikation von Pufulete 2022 in der Indikation koronare Herzkrankheiten und von den externen Sachverständigen in der Indikation RRMS) wurde im Arbeitspapier das Vorgehen in Bezug auf mögliche Optimierungsansätze und Variationsmöglichkeiten diskutiert und ob beispielsweise einer der in Pufulete 2022 vorgesehenen Bausteine (systematische Literaturrecherche, Interviews mit Klinikerinnen und Klinikern und Umfrage unter Klinikerinnen und Klinikern) in Abwägung mit dem dafür notwendigen Ressourcenaufwand verzichtbar erscheint.

4.1.3 Überprüfung der Machbarkeit einer Confounderidentifikation gemäß Pufulete 2022

Wenn die externe Gruppe den Bericht inklusive der identifizierten Confounder bereitstellte und die Prüfung durch das IQWiG ergab, dass adäquat gemäß den Vorgaben in Pufulete 2022 vorgegangen wurde, sollte der in Pufulete 2022 beschriebene Ansatz als grundsätzlich machbar eingeschätzt werden. Es sollte zudem beurteilt werden, inwieweit der im Rahmen der Vergabe angestrebte Zeitrahmen eingehalten werden konnte.

4.2 Eigene Untersuchungen

Basierend auf dem Bericht der externen Sachverständigen wurden weitere eigene Untersuchungen durchgeführt, die über den methodischen Ansatz von Pufulete 2022 hinausgehen. Diese werden im Folgenden beschrieben.

4.2.1 Zusammenführung der identifizierten potenziellen Confounder

Es wurde geprüft, ob die im Bericht der externen Sachverständigen identifizierten potenziellen Confounder-Variablen weitergehend zusammengeführt werden können. Dazu wurde von 2 Personen unabhängig voneinander geprüft, welche Confounder-Variablen ein inhaltlich gleiches Konstrukt messen oder sich inhaltlich überlappen. Es wurde dokumentiert, welche einzelnen Confounder-Variablen zusammengefasst werden. Anschließend wurden für die zusammengeführten Confounder geeignete Operationalisierungen definiert.

Alle Schritte wurden in der Projektgruppe konsentiert. Diskrepanzen wurden durch Diskussion aufgelöst. Darüber hinaus wurde ein Workshop mit den externen Sachverständigen durchgeführt, in dem die Liste der potenziellen Confounder und deren Operationalisierungen diskutiert und anschließend auf Basis der Expertise der externen Sachverständigen und der internen Projektgruppe überarbeitet wurde.

4.2.2 Fokussierte Informationsbeschaffung zu Publikationen zur systematischen Vorgehensweise der Confounderidentifikation

Anhand einer fokussierten Informationsbeschaffung wurde geprüft, ob neben dem von Pufulete 2022 vorgeschlagenen systematischen Ansatz zur Confounderidentifikation weitere Publikationen zur systematischen Confounderidentifikation vorliegen.

Eingeschlossen wurden Publikationen, in denen vergleichbar zum Vorgehen in Pufulete 2022, ein systematisches Vorgehen zur Confounderidentifikation mit dem Ziel a priori ein Confounder-Set zu definieren, beschrieben wurde. Publikationen mit datengestützten oder datenbasierten Ansätzen der Identifikation (bspw. statistische Verfahren basierend auf Daten aus elektronischen Patientenakten [25,26]) wurden ausgeschlossen. Es erfolgte keine Einschränkung des Publikationszeitraums. Die Publikation musste in deutscher oder englischer Sprache verfasst sein.

Folgende Informationsquelle sowie Suchtechnik wurde für die fokussierte Informationsbeschaffung berücksichtigt: MEDLINE und die „Similar Articles“-Funktion in Pubmed zu vorab bekannten Publikationen [2,11,27,28].

Die identifizierten Treffer wurden in einem 1. Schritt anhand ihres Titels und, sofern vorhanden, Abstracts in Bezug auf ihre potenzielle Relevanz bezüglich der Einschlusskriterien (siehe oben, Kriterien für den Einschluss) bewertet. Als potenziell relevant erachtete Dokumente wurden in einem 2. Schritt anhand ihres Volltextes auf Relevanz geprüft. Beide Schritte der Selektion erfolgten durch 1 Person.

4.3 Ableitung von Anforderungen an eine systematische Confounderidentifikation (Handlungsempfehlungen)

Anhand der Publikation von Pufulete 2022, der Erfahrungen aus dem Bericht der externen Sachverständigen zur Identifikation von Confoundern gemäß Pufulete 2022 und anhand der weiterführenden eigenen Untersuchungen wurden konkrete Handlungsempfehlungen zur systematischen Identifikation von Confoundern für nicht randomisierte vergleichende Studien abgeleitet. Hierbei wurde im Speziellen geprüft, inwieweit sich für Fragestellungen mit erwartbar weniger umfassender Studienlage (wie beispielsweise Fragestellungen der anwendungsbegleitenden Datenerhebung) besondere Anforderungen ergeben.

5 Ergebnisse

5.1 Confounderidentifikation gemäß Pufulete 2022 für die Indikation RRMS

5.1.1 Ergebnis der Confounderidentifikation nach Pufulete 2022

Das Ziel des externen Berichts war es, systematisch und unabhängig potenzielle Confounder zu identifizieren, die bei der Konzeption und der Durchführung eines nicht randomisierten Vergleichs der Interventionen Dimethylfumarat vs. Glatirameracetat hinsichtlich patientenrelevanter Endpunkte bei Erwachsenen in der Indikation RRMS berücksichtigt werden sollten. Das Ergebnis des Berichts der externen Sachverständigen wird im Folgenden kurz zusammengefasst. Anschließend wird das methodische Vorgehen im Abgleich mit Pufulete 2022 [11] diskutiert. Der vollständige Bericht der externen Sachverständigen ist in Anhang D dargestellt.

Die externen Sachverständigen haben über den systematischen 3-stufigen Ansatz der Confounderidentifikation 160 potenzielle Confounder identifiziert. 125 potenzielle Confounder wurden dabei über die Literaturrecherche (Suche nach Beobachtungsstudien und RCTs) identifiziert. Über die Interviews (n = 8) mit Klinikerinnen und Klinikern aus Deutschland und aus der Schweiz wurden insgesamt 53 potenzielle Confounder identifiziert, 18 davon wurden auch durch die Literaturrecherche ermittelt, 35 wurden erstmals durch die Klinikerinnen und Kliniker genannt. Im anschließenden quantitativen Survey mit 21 Personen (von denen 7 die Umfrage beendeten) wurden 137 der 160 identifizierten Confounder dahin gehend beurteilt, wie stark diese die Therapieentscheidung der Ärztin oder des Arztes beeinflussen. Die Einstufung erfolgte mittels einer Skala von "1 – keine Berücksichtigung" bis "5 – sehr starke Berücksichtigung". 40 der 137 Confounder wurden im Mittel mit einem Wert > 4 bewertet.

Die externen Sachverständigen haben 148 (93 %) der 160 potenziellen Confounder als „wahrscheinlich ein Prädiktor für die Behandlung“ und von diesen 148 potenziellen Confoundern 128 (80 % der Gesamtzahl) als „wahrscheinlich auch ein Prädiktor für die betrachteten Endpunkte“ eingestuft. 11 (7 %) potenzielle Confounder wurden als unklar eingestuft, was die externen Sachverständigen als „wahrscheinlich weder ein Prädiktor für die Behandlung noch für die Endpunkte“ interpretieren. Für den potenziellen Confounder Type of MS wurde keine Einstufung durch die externen Sachverständigen vorgenommen, da nur die RRMS untersucht wurde.

In einem letzten Schritt wurden die 160 potenziellen Confounder von den externen Sachverständigen zu 136 Confoundern geclustert (d. h. Confounder, die auf ähnliche Weise gemessen oder bewertet werden können).

5.1.2 Diskussion des Vorgehens der externen Sachverständigen

Die externen Sachverständigen legen einen ausführlichen Bericht mit einer umfassenden Confounder-Recherche nach Pufulete 2022 vor (siehe Anhang D). Die übergeordneten Ergebnisse der einzelnen Schritte wurden detailliert im Bericht dokumentiert. Eine detaillierte Dokumentation der Datenquelle jedes einzelnen Confounders (inkl. wörtlichem Zitat) liegt als separate Excel-Datei (Anhang F im Bericht der externen Sachverständigen [Datensatz] vor [29]).

Insgesamt bearbeiteten die externen Sachverständigen die Fragestellung angemessen entsprechend den Vorgaben in Pufulete 2022. Im Folgenden wird detailliert auf Abweichungen zu dem Vorgehen gemäß Pufulete 2022 und ggf. vorgenommene Anpassungen, z. B. zur Aufwandsreduzierung, eingegangen und es werden mögliche Optimierungsansätze bzw. Variationsmöglichkeiten beschrieben.

Systematische Literaturrecherche

Ein- und Ausschlusskriterien und Selektion

Die externen Sachverständigen berücksichtigen in ihrer Literaturrecherche zusätzlich zu RCTs und Kohortenstudien auch explizit Non-RCTs (gemäß der Taxonomie von Viswanathan [30]). Dort sind Non-RCTs definiert als prospektive Studien, bei der Einzelpersonen oder Gruppen durch eine nicht zufällige Methode (z. B. Geburtsdatum, Aufnahmedatum, Urteil des Prüfers) der Intervention oder der Kontrolle zugewiesen werden. Der Einschluss von Non-RCTs geht über den Ansatz von Pufulete 2022 hinaus. Allerdings sind entsprechend der Angaben in Tabelle 2 des Berichts der externen Sachverständigen die genannten Confounder lediglich aus Beobachtungsstudien oder randomisierten Studien identifiziert worden. Da nach der oben aufgeführten Definition Non-RCTs keine Beobachtungsstudien sind, geht aus der zusätzlichen Berücksichtigung von Non-RCTs im vorliegenden Fall kein zusätzlicher Erkenntnisgewinn hervor.

In den Einschlusskriterien und bei der Selektion der relevanten Studien haben die externen Sachverständigen neben der Population und den Interventionen auch die zu untersuchenden Endpunkte berücksichtigt. Darüber hinaus mussten für den Studieneinschluss 90 % der Studienteilnehmer der Indikation entsprechen. Diese im Vergleich zu Pufulete 2022 etwas differenzierteren Einschlusskriterien sind nachvollziehbar, insbesondere um sicherzustellen, dass die Studie die Studienpopulation der Fragestellung in ausreichendem Umfang abbildet.

Extraktion und Sättigungskriterium

Analog zum Vorgehen von Pufulete 2022 extrahierten die externen Sachverständigen neben Studien- und Patientencharakteristika auch die Interventionen und alle Faktoren, die zur Adjustierung in den statistischen Analysen verwendet wurden. Bei RCTs wurden die in tabellarischer Form berichteten Patientencharakteristika („Tabelle 1“) extrahiert. Darüber

hinaus extrahierten die externen Sachverständigen den prozentualen Anteil an untersuchten Frauen und Mittelwert und Standardabweichung des Alters zu Studienbeginn und vermerkten für alle identifizierten potenziellen Confounder, für welchen Endpunkt sie in der Analyse verwendet wurden. Insbesondere die zusätzliche Extraktion von Alter und Geschlecht erscheint im Sinne der Einsparung von Ressourcen verzichtbar, zumal die externen Sachverständigen keine Konsequenzen aus diesen Informationen gezogen haben.

Pufulete 2022 verwendeten bei der Extraktion der Confounder in der Literaturrecherche ein Sättigungskriterium. Wenn innerhalb von 10 konsekutiven Studien kein neuer potenzieller Confounder identifiziert wurde, war die Confounder-Extraktion beendet. Die externen Sachverständigen wendeten ein Sättigungskriterium nur für Studien vor 2020 an. Diese wurden zufällig ausgewählt und die Extraktion mithilfe eines Sättigungskriteriums gestoppt, wenn innerhalb von 15 konsekutiven Studien kein neuer potenzieller Confounder identifiziert wurde. Von den relevanten Studien, die seit 2020 publiziert wurden (n = 40 der 351 aus der Suche identifizierten), wurden hingegen alle identifizierten Confounder extrahiert. Die externen Sachverständigen begründen diese Unterteilung in Abhängigkeit des Zeitpunkts der Veröffentlichung der Studien damit, dass sich das Gebiet der multiplen Sklerose im Vergleich zu Herz-Kreislauf-Erkrankungen (Gegenstand der Untersuchungen bei Pufulete 2022), bei denen die Krankheitsmechanismen seit Jahrzehnten relativ gut bekannt sind, viel schneller entwickle. Um insbesondere diese Entwicklungen bei multipler Sklerose durch neuere Literatur und neue Quellen von Beobachtungsdaten angemessen zu berücksichtigen, unterteilten die externen Sachverständigen die Studien basierend auf dem Zeitpunkt der Veröffentlichung. Die Begründung der externen Sachverständigen, die Recherche hinsichtlich des Zeitpunkts der Veröffentlichung zu unterteilen, ist nachvollziehbar. Aus dem Bericht der externen Sachverständigen geht jedoch nicht hervor, wie viele zusätzliche potenzielle Confounder durch den Verzicht auf ein Sättigungskriterium (für Studien nach 2020) bzw. die Verwendung von 15 statt 10 konsekutiven Studien als Sättigungskriterium (für Studien vor 2020) identifiziert wurden. Daher kann anhand der vorliegenden Informationen im Bericht der externen Sachverständigen nicht eingeschätzt werden, inwiefern der Mehraufwand den zusätzlichen Gewinn an Confoundern (57/125 [46 %] wurden ausschließlich in Studien identifiziert, die 2020 oder später veröffentlicht wurden) rechtfertigt. Die externen Sachverständigen gehen davon aus, eine vollständigere Liste identifiziert zu haben, aber beschreiben den gewählten Ansatz selbst als sehr ressourcenintensiv.

Die Volltexte wurden von den externen Sachverständigen zusätzlich nach den Schlagworten „mediator“, „modifier“, „causal“, „mediate“, „mediating“, „acyclic“ und „dag“ durchsucht. Diese Schlagwort-Suche war in der vorliegenden Indikation RRMS nicht erfolgreich („causal“ wurde zwar 3-mal genannt, jedoch in inhaltlich anderem Zusammenhang, „acyclic“ nur einmal [siehe Tabelle 3 im Bericht der externen Sachverständigen]). Trotzdem erscheint dieser Ansatz plausibel. Um diesen Schritt zu vereinfachen, kann die Schlagwortsuche durch die einfache

Durchsicht der Publikation nach einem gerichteten azyklischen Graphen (DAG) ersetzt werden. Wird ein DAG identifiziert, sollten die enthaltenen Variablen als potenzielle Confounder-Variablen extrahiert werden (siehe auch Abschnitt 5.3).

Qualitative Interviews mit klinischen Expertinnen und Experten

Die Einbindung von klinischen Expertinnen und Experten stellt bei Pufulete 2022 den 2. Baustein der systematischen Confounderidentifikation dar. Sie orientieren sich hierbei an den Cochrane-Empfehlungen, die auch in das Instrument ROBINS-I zur Bewertung von Verzerrungen nicht randomisierter Studien übernommen wurden [12,13].

Auch die externen Sachverständigen haben halbstrukturierte Interviews mit klinischen Expertinnen und Experten durchgeführt (n = 8). Die externen Sachverständigen entwickelten analog zu Pufulete 2022 einen vignettenbasierten Interviewleitfaden (mit 5 Szenarien zu klinischen Fällen, siehe Anhang B im Bericht der externen Sachverständigen). Die klinischen Expertinnen und Experten wurden anhand ihrer Erfahrung in der Verordnung der zu untersuchenden Interventionen (Dimethylfumarat oder Glatirameracetat) bei Patientinnen und Patienten mit multipler Sklerose ausgewählt.

Über die Interviews konnten 35 weitere potenzielle Confounder (35/160 [22 %] der potenziellen Confounder) identifiziert werden, die über die systematische Literaturrecherche nicht gefunden wurden. Dazu gehörten insbesondere patientenzentrierte Faktoren (z. B. der Patientenwunsch oder das Sicherheitsbedürfnis). Diese Beobachtungen entsprechen dem Ergebnis von Pufulete 2022.

Pufulete 2022 gehen in ihrer Publikation davon aus, dass mit 12 Expertinnen und Experten (je 6 in Kardiologie und Herzchirurgie) eine Sättigung erreicht werden kann. Die externen Sachverständigen planten ursprünglich 10 Teilnehmende zu akquirieren. Aufgrund von Schwierigkeiten bei der Rekrutierung wurden letztlich nur 8 klinische Expertinnen und Experten eingebunden. Die externen Sachverständigen geben an, dass die Beiträge der 8 Expertinnen und Experten gesättigt waren und die bei der Therapieentscheidung berücksichtigten Faktoren ausreichend widerspiegeln. Die Einbindung von 8 klinischen Expertinnen und Experten für die qualitativen Interviews wird insgesamt als angemessen beurteilt. Ob ggf. auch eine geringere Zahl an Interviews bereits eine Sättigung erzielt hätte, kann an dieser Stelle nicht beurteilt werden.

Dem Bericht der externen Sachverständigen ist lediglich zu entnehmen, ob eine Confounder-Variable auch im Interview genannt wurde, nicht aber die Anzahl der Nennungen. Dies ist angemessen, da es in diesem Schritt allein um die Identifikation neuer potenzieller Confounder geht. Aus der Anzahl der klinischen Expertinnen und Experten, welche den potenziellen Confounder benennen, lassen sich keine Rückschlüsse auf z. B. die Wichtigkeit des potenziellen Confounders ziehen.

Quantitative Umfrage mit klinischen Expertinnen und Experten

Als 3. Baustein wurde von den externen Sachverständigen analog zum Vorgehen von Pufulete 2022 eine quantitative Umfrage unter klinischen Expertinnen und Experten durchgeführt. Hierzu wurden die aus der Literaturrecherche und den Interviews identifizierten Confounder zusammengeführt. Die externen Sachverständigen haben darüber hinaus für die Umfrage die 160 identifizierten Confounder in ersten Schritten zusammengefasst und reduziert (13 über ähnliche Confounder abgedeckt, 2 ohne Studienkontext nicht beurteilbar, 8 Confounder waren zu Beginn der Umfrage noch nicht identifiziert). Schließlich wurden den Teilnehmerinnen und Teilnehmern der Umfrage 137 Confounder zur Bewertung vorgelegt. Abweichend zu Pufulete 2022 sollten die Teilnehmenden der Umfrage nicht die 5 wichtigsten Confounder benennen, sondern auf einer 5-Punkte-Skala mit einem Score von 1 – keine Berücksichtigung bis 5 – starke Berücksichtigung bewerten, wie stark jeder einzelne der 137 potenziellen Confounder ihre Entscheidung bei der Verschreibung von Dimethylfumarat oder Glatirameracetat beeinflusst. Zwar hatten die externen Sachverständigen geplant, ähnlich wie Pufulete 2022, die klinischen Expertinnen und Experten nach den 10 wichtigsten Confoundern zu fragen, haben auf diesen Schritt aber schließlich verzichtet, um die Durchführbarkeit der Umfrage zu verbessern und weitere Hürden bei der Rekrutierung zu vermeiden. Das Vorgehen der externen Sachverständigen gab den klinischen Expertinnen und Experten zusätzlich die Möglichkeit weitere potenzielle Confounder über ein Freitextfeld anzugeben.

Die externen Sachverständigen haben 79 Personen zur Umfrage eingeladen, 21 Personen haben die Umfrage begonnen und 7 Personen haben die Umfrage beendet. Die niedrigen Rücklaufquoten könnten auf die hohe Anzahl der zu bewertenden potenziellen Confounder und dem damit verbundenen Ressourceneinsatz für die klinischen Expertinnen und Experten zurückzuführen sein. Dies vermuteten auch die externen Sachverständigen. Welche Rücklaufquoten bei Pufulete 2022 erreicht werden konnten, geht aus der Publikation nicht hervor. Es ist lediglich angegeben, dass insgesamt 110 Klinker die Umfrage beendeten.

Das geplante Vorgehen der externen Sachverständigen, anstatt der Nennung der 10 wichtigsten Confounder ein Scoring der einzelnen Confounder durch die klinischen Expertinnen und Experten durchführen zu lassen, stellt eine alternative Herangehensweise dar, die Relevanz der Confounder einzuschätzen. Dieses Vorgehen ist durchaus nachvollziehbar. Auch wenn von den beiden möglichen Bewertungsmethoden das Scoring einen höheren Aufwand bedeutet, ist die Bewertung eines jeden einzelnen Confounders im Vergleich zu dem von Pufulete 2022 gewählten Vorgehens der Benennung der 5 wichtigsten Confounder in der Gesamtschau inhaltlich zu befürworten. Mithilfe der Scores können alle potenziellen Confounder hinsichtlich ihrer Bedeutung sortiert werden. Diese Rangfolge kann nachgelagert ggf. als Entscheidungshilfe genutzt werden (z. B. in Fällen in denen das statistische Modell nicht konvergiert kann auf Basis der Scores begründet werden, welcher potenzielle Confounder bei der Adjustierung im Modell als erstes verzichtbar erscheint). Um

den Aufwand für diese Confounder-Bewertung zu reduzieren, wird in Abschnitt 5.3 ein möglicher Ansatz der intensiveren Zusammenführung der Liste der zu bewertenden potenziellen Confounder vorgeschlagen.

Über die Möglichkeit, weitere potenzielle Confounder über ein Freitextfeld anzugeben, wurden keine weiteren potenziellen Confounder identifiziert. In der von Pufulete 2022 beschriebenen Methodik ist diese Möglichkeit nicht vorgesehen, die externen Sachverständigen geben keine Begründung für diesen zusätzlichen Schritt. Aufgrund des erwartbar geringen Informationsgewinns ist dieser Schritt verzichtbar.

Kategorisierung und Klassifikation

In einem 4. und letzten Schritt wurde von den externen Sachverständigen beurteilt, ob der potenzielle Confounder sowohl die Therapieentscheidung als auch das Behandlungsergebnis beeinflusst. Für 148 der 160 potenziellen Confounder (93 %) wurde es als wahrscheinlich eingestuft, dass sie ein Prädiktor für die Therapieentscheidung sind. 128 (80 %) potenzielle Confounder beeinflussen wahrscheinlich auch das Behandlungsergebnis. 11 (7 %) potenzielle Confounder wurden als unklar eingestuft, was die externen Sachverständigen als „wahrscheinlich weder ein Prädiktor für die Behandlung noch für die Endpunkte“ interpretieren. Für den potenziellen Confounder Type of MS wurde keine Einstufung durch die externen Sachverständigen vorgenommen, da nur die RRMS untersucht wurde. Ähnlich hierzu haben Pufulete 2022 in einem letzten Schritt eine Klassifizierung der potenziellen Confounder-Variablen danach vorgenommen, ob sie Prädiktor sind sowohl für die Behandlung als auch den Endpunkt, oder nur für eine von beiden oder für keines. Pufulete 2022 stuften 34 der 70 identifizierten potenziellen Confounder (49 %) als Prädiktor sowohl für die Behandlung als auch den Endpunkt ein.

Das Ziel einer solchen Kategorisierung ist, „wahre Confounder“ zu identifizieren. Da die Einschätzung, in welchem Zusammenhang die Variablen stehen, auf starken Annahmen beruht und entsprechend mit Unsicherheiten behaftet ist, ist eine solche Identifikation der Confounder in der Regel nicht möglich. Entsprechend sollten Confounder-Variablen, deren Einfluss auf Therapieentscheidung und Behandlungsergebnis unsicher ist, nicht ausgeschlossen werden. Dieser Schritt der Kategorisierung für jede einzelne Variable ist entsprechend verzichtbar. Lediglich eine literaturgestützte Begründung kann den Ausschluss eines potenziellen Confounders gegebenenfalls rechtfertigen (siehe Abschnitt 5.3).

Zudem wurden in diesem letzten Schritt die 160 potenziellen Confounder von den externen Sachverständigen zu 136 potenziellen Confoundern geclustert und verschiedenen Kategorien, wie z. B. Komorbiditäten oder klinischer Status, zugeordnet. Das Cluster Lifestyle besteht beispielsweise aus den potenziellen Confoundern Tabakkonsum (tobacco use), Art der beruflichen Tätigkeit (type of professional work activity) und Größe und Gewicht (height and

weight). Unter letzterem wurden die 3 potenziellen Confounder Body-Mass-Index, Übergewicht oder Adipositas (overweight or obesity) und Gewicht (weight) zusammengefasst. Eine solche Clusterung wurde von Pufulete 2022 nicht durchgeführt, aber in der Diskussion der Publikation als sinnvoll beschrieben, da es sich hierbei um hoch korrelierte Variablen handle und die Clusterung die Zahl der in der statistischen Analyse zu berücksichtigenden Confounder reduzieren könne. Das Ergebnis eigener Untersuchungen, welche Möglichkeiten der weitergehenden Zusammenführung der von den externen Sachverständigen identifizierten potenziellen Confounder bestehen, werden in Abschnitt 5.2.1 beschrieben.

5.1.3 Machbarkeit einer Confounderidentifikation gemäß Pufulete 2022

Die Prüfung des Berichts der externen Sachverständigen ergab, dass im Wesentlichen adäquat gemäß den Vorgaben in Pufulete 2022 vorgegangen wurde. Die externen Sachverständigen konnten mithilfe der systematischen Literaturrecherche und den Interviews mit klinischen Expertinnen und Experten insgesamt 160 potenzielle Confounder identifizieren. Grundsätzlich wird die Durchführung der Confounderidentifikation nach dem Ansatz von Pufulete 2022 somit als machbar bewertet.

Die externen Sachverständigen haben für die Identifikation der 160 Confounder 62 Personentage (à 8 Arbeitsstunden) benötigt (siehe Tabelle 6 im Bericht der externen Sachverständigen). Die Projektplanung, das Projektmanagement, die Supervision und die Erstellung des Berichtes wurden in dem geschätzten Zeitaufwand nicht berücksichtigt.

Der Zeitrahmen für diese Projektbestandteile, die als Teil des gesamten Identifikationsprozesses angesehen werden, ist somit unklar. Die externen Sachverständigen selbst beschreiben in der Diskussion des Berichts, dass die methodische und klinische Expertise ihres Teams mehrere Schritte erleichtert haben könnte und damit der angegebene Zeitaufwand ggf. unterschätzt ist.

Zur Überprüfung der Machbarkeit sollte auch eine Beurteilung erfolgen, inwieweit der im Rahmen der Vergabe angestrebte Zeitrahmen eingehalten werden konnte. Aufgrund abgestimmter, nachträglicher Anpassungen des Vorgehens, ohne erneute Aktualisierung der angestrebten Zeiten, erscheint ein Abgleich der benötigten Zeit mit der im initialen Angebot angegebenen Zeit nicht zielführend. Daher wird im Folgenden primär Bezug auf den im Bericht geschätzten Zeitaufwand genommen.

Die systematische Literaturanalyse stellt mit 33 geschätzten Personentagen den zeitaufwendigsten Baustein der Confounderidentifikation dar. Hierbei ist zu berücksichtigen, dass die Indikation RRMS unter anderem aus dem Grund ausgewählt wurde, dass insgesamt von einer ausreichenden Studienlage ausgegangen wird. So ergab die Recherche nach Abzug von Duplikaten 1202 Treffer, die gescreent werden mussten. Daraus wurden 80 Studien identifiziert, bei denen die entsprechenden Informationen extrahiert werden mussten. Dies

ist in einer ähnlichen Größenordnung wie bei Pufulete 2022. Pufulete 2022 screenen 2544 Publikationen ohne Duplikate und extrahierten schließlich aus 47 geeigneten Publikationen. In Indikationen mit erwartbar weniger umfassender Studienlage (und entsprechend einer deutlich geringeren Anzahl an zu screenenden und zu extrahierenden Treffern der Literaturrecherche) kann der Zeitaufwand für diesen Baustein deutlich geringer ausfallen. Darüber hinaus ist der von den externen Sachverständigen gewählte Ansatz, für Publikationen nach 2020 kein Sättigungskriterium für die Confounderidentifikation zu wählen, sehr sensitiv. In diesem Schritt besteht somit ggf. Einsparpotenzial durch die Verwendung eines Sättigungskriteriums für alle Publikationen.

Für den Baustein der qualitativen Interviews geben die externen Sachverständigen einen Zeitbedarf von 16 Personentagen an. Es wurden 8 Interviews durchgeführt. Es ist unklar, wie lange die Interviews gedauert haben, Pufulete 2022 geben die Dauer pro Interview mit 26 bis 45 Minuten an. Der zeitaufwendigste Teil dieses Bausteins wird entsprechend die Erarbeitung des Interviewleitfadens ebenso wie die Transkription und Auswertung der Interviews sein.

Der quantitative Survey und die Kategorisierung und Klassifizierung machen mit 8 bzw. 5 Personentagen den kleinsten Zeitbedarf aus. Auch wenn die beiden Bausteine für die externen Sachverständigen den geringsten Zeitaufwand bedeuten, ist der Ressourceneinsatz für die eingebundenen klinischen Expertinnen und Experten hoch. Dieser Zeitaufwand kann durch die Zusammenführung der Confounder-Variablen vor dem quantitativen Survey deutlich reduziert werden (siehe Abschnitt 5.3.2).

5.2 Eigene Untersuchungen

5.2.1 Potenzielle Confounder für die Indikation RRMS

Die externen Sachverständigen identifizierten 160 potenzielle Confounder-Variablen und clusterten Confounder, die auf ähnliche Weise gemessen oder bewertet werden können zu 136 potenziellen Confoundern. Diese geringe Reduzierung verdeutliche laut den externen Sachverständigen, dass es unter diesen mehr als 100 Faktoren tatsächlich zahlreiche eindeutig unterschiedliche konzeptionelle Aspekte gebe, die berücksichtigt werden müssen, um „unverzerrte Schätzungen“ der Auswirkungen von Behandlungsentscheidungen zu erhalten.

Eine Zusammenführung der über die Literaturrecherche und die Interviews mit klinischen Expertinnen und Experten identifizierten Confounder war nicht Bestandteil der Ausarbeitung von Pufulete 2022, wird von den Autorinnen und Autoren aber als logischen weiteren Schritt diskutiert. Insgesamt gibt es für diesen Schritt aber bislang keine Handlungsempfehlungen.

Mit der Zielsetzung, eine randomisierte Studie anhand von Registerdaten nachzubilden, erscheint die vollständige Erfassung von 160 bzw. 136 Confoundern nicht sinnvoll. Zum einen bedeutet das, dass alle diese Confounder in der Datenquelle, bspw. einem Register, für alle in

die Studie eingeschlossenen Patientinnen und Patienten mit hoher Datenqualität erhoben werden müssen. Zum anderen wird ein solcher Umfang an Adjustierungsvariablen auch in Abhängigkeit von der zur Verfügung stehenden Patientenzahl im zu untersuchenden Datensatz absehbar zu Problemen in den statistischen Analysen führen [31,32]. Die in Tabelle 5 des Berichts der externen Sachverständigen präsentierte Liste der 136 geclusterten Confounder zeigt an verschiedenen Stellen trotz der Clusterung von Confoundern (die auf ähnliche Weise gemessen oder bewertet werden), inhaltlich überlappende Konzepte (z. B. Land und Region; gesundheitsbezogene Lebensqualität indikationsspezifisch und generisch).

Aus diesem Grund wurde überprüft, inwieweit eine weitergehende intensivere Zusammenführung der identifizierten potenziellen Confounder-Variablen vorgenommen werden kann, die höchstens mit einem vertretbaren Informationsverlust hinsichtlich der identifizierten potenziellen Confounder einhergeht. Während der Bearbeitung wurde außerdem deutlich, dass einige der identifizierten Confounder-Variablen basierend auf der vorliegenden Fragestellung zur Nachbildung der RCT CONFIRM ausgeschlossen werden können (siehe nachfolgender Abschnitt). Die vorgenommenen Schritte werden im Folgenden beschrieben und sind in Abbildung 2 zusammenfassend dargestellt.

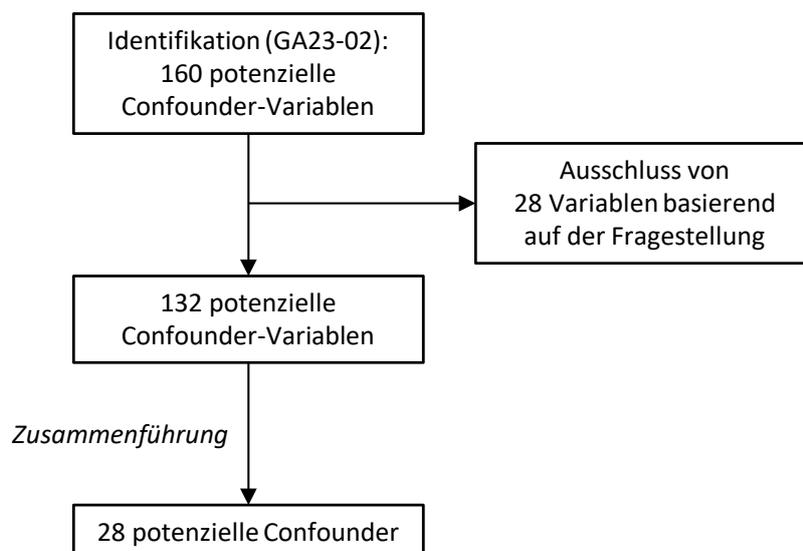


Abbildung 2: Ergebnis der Zusammenführung der potenziellen Confounder-Variablen

Im Folgenden wird im Sinne einer einheitlichen Benennung der Begriff Confounder-Variable für die im 1. Schritt extrahierten Variablen verwendet. Der Begriff Confounder bezeichnet das Ergebnis der zusammengeführten Confounder-Variablen.

Das Ergebnis der Zusammenführung in Form einer Confounder-Liste inklusive Vorschläge für Operationalisierungen wurde den externen Sachverständigen in einem Workshop vorgestellt und mit ihnen diskutiert (siehe Abschnitt unten).

Ausschluss von Variablen basierend auf der vorliegenden Fragestellung

Die vorliegende systematische Confounderidentifikation soll zur Durchführung eines nicht randomisierten Vergleichs der Interventionen Dimethylfumarat vs. Glatirameracetat bei Patientinnen und Patienten ab 18 Jahren in der Indikation RRMS genutzt werden können. Im Folgenden wird angenommen, dass die RCT CONFIRM [24] als Modellstudie mittels Target Trial Emulation als nicht randomisierte vergleichende Studie nachgebildet werden soll. Somit konnten die konkreten Ein- und Ausschlusskriterien der RCT CONFIRM genutzt werden, um Variablen, die für die vorliegende Fragestellung keine Relevanz haben, auszuschließen. Darüber hinaus wurden weitere methodische Gründe identifiziert, die im Sinne der Einhaltung der methodischen Grundsätze einer Target Trial Emulation keine Relevanz haben. Insgesamt wurden 28 Variablen ausgeschlossen. Die Ausschlussgründe lassen sich folgendermaßen zusammenfassen:

- Ausschluss basierend auf fehlender Relevanz im deutschen Versorgungskontext (n = 4)
- Ausschluss basierend auf der Studie CONFIRM (anhand der Ein-/Ausschlusskriterien im Studienprotokoll [24]) (n = 14)
- Ausschluss basierend auf methodischen Gründen (z. B. Studienverlaufscharakteristikum) (n = 4)
- Ausschluss, da Originalpublikation außerhalb der Fragestellung (n = 6)

Im Folgenden werden zur besseren Nachvollziehbarkeit Beispiele der ausgeschlossenen Variablen beschrieben. Alle 28 identifizierten Variablen, die basierend auf einem der Ausschlussgründe nicht weiter betrachtet wurden, sind in Anhang B, Tabelle 3 aufgelistet.

Variablen, wie der Versicherungsstatus (health insurance status) oder die Kosten für Krankenhausbehandlungen (z. B. previous hospital care costs) wurden aufgrund der fehlenden Relevanz im deutschen Versorgungskontext ausgeschlossen.

Insgesamt wurden 14 Variablen basierend auf der fehlenden Relevanz für die Nachbildung der RCT CONFIRM nicht weiter berücksichtigt. Dies sind zum einen Variablen, die entsprechend der Einschlusskriterien für alle Patientinnen und Patienten die gleiche Ausprägung annehmen müssen. Ein Beispiel hierfür ist das McDonald-Diagnosekriterium zur Klassifikation der multiplen Sklerose (McDonald criteria, MS classification), das alle Patientinnen und Patienten erfüllen müssen. Zum anderen handelt es sich um Variablen, die aufgrund der Ausschlusskriterien für keine Patientin bzw. keinen Patienten zutreffen. Beispielsweise ist das Vorliegen einer Hepatitis ein Ausschlusskriterium der Studie CONFIRM. Es ist jedoch zu beachten, dass bei spezifischen Ausschlusskriterien oder Kontraindikationen des Arzneimittels geprüft werden muss, ob der Confounder nicht weitere Ausprägungen umfasst, die einem Ausschluss widersprechen. Zum Beispiel ist es nicht sachgerecht, den Confounder absolute

Lymphozytenzahl, Lymphopenie (absolute Lymphocyte count, Lymphopenia, stetige Operationalisierung), basierend auf der Kontraindikation schwere Lymphopenie von Dimethylfumarat auszuschließen. Dies kann allenfalls nach Einschränkung der Operationalisierung des Confounders auf schwere Lymphopenie erfolgen.

Zwei der 4 Variablen (Time on treatment und Time on study), die aufgrund methodischer Gründe ausgeschlossen wurden, sind zeitabhängig und stellen Studienverlaufscharakteristika dar. Für die vorliegende Fragestellung wird aber davon ausgegangen, dass lediglich Baselinewerte als Confounder berücksichtigt werden. Daher wurden auch die 2 Variablen (time since last visit und Expanded Disability Status Scale (EDSS) scores/improvement at the end of the study treatment/last check-up) ausgeschlossen.

Für weitere 6 Variablen lag die berücksichtigte Originalpublikation außerhalb der vorliegenden Fragestellung. Zum Beispiel wurden aus einer Publikation zur Untersuchung schwangerschaftsbedingter Schübe folgende 3 Variablen extrahiert: Disease modifying therapy (DMT) reinitiation postpartum, Number of visits 1 year postpartum, Number of visits in pregnancy. Eine bestehende Schwangerschaft war ein Ausschlusskriterium der Studie CONFIRM. Entsprechend wird dieses spezifische Patientenkollektiv basierend auf der vorliegenden Fragestellung, die Studie CONFIRM nachzubilden, nicht weiter berücksichtigt.

Zusammenführung

Die 132 verbleibenden potenziellen Confounder-Variablen wurden mit dem Ziel, Duplikate und Überlappungen zu reduzieren, zusammengeführt. Damit sollte höchstens ein vertretbarer Informationsverlust hinsichtlich der identifizierten potenziellen Confounder einhergehen. Hierzu wurden gleiche oder inhaltlich überlappende Confounder-Variablen zusammengeführt. Insgesamt konnten die 132 potenziellen Confounder-Variablen zu 28 potenziellen Confoundern verdichtet werden. Das Set der 28 potenziellen Confounder inklusive der vorgeschlagenen Operationalisierungen ist nachfolgend in Tabelle 1 dargestellt. Auf die Auswahl der Operationalisierungen wird im nachfolgenden Abschnitt eingegangen. Die Zuordnung, welche potenziellen Confounder-Variablen unter dem jeweiligen potenziellen Confounder zusammengefasst wurden, ist in Anhang B, Tabelle 2 aufgeführt.

Beispielsweise wurden dem Confounder gesundheitsbezogene Lebensqualität (Quality of Life) die potenziellen Confounder-Variablen EQ-5D, SF-36 Scores, Multiple Sclerosis Impact Scale, Multiple Sclerosis International Questionnaire of Quality of Life und Sexual Satisfaction Scale Score zugeordnet. Unter dem Confounder Komorbiditäten wurden folgende potenzielle Confounder-Variablen zusammengefasst: Comorbidities (general), Depression diagnosis, Hamilton Anxiety Rating Scale (HAM-A), Hamilton Depression Rating Scale (HAM-D), allergy, cardiovascular comorbidities, Charlson Comorbidity Index (CCI) score, chronic infections, dermatological diseases, diabetes, headache, skin problems, thyroid dysfunction,

tuberculosis, bowel symptoms, Depression (self-rated), gastrointestinal disorders, hospitalisations in previous year.

Für einige der identifizierten potenziellen Confounder-Variablen ist eine Zuordnung zu verschiedenen Confoundern möglich (z. B. wurde die Confounder-Variable Magnetic resonance imaging [MRI] status dem Confounder Number of T2 lesions zugeordnet). Die Zuordnung zu einem der potenziellen Confoundern ist ausreichend, da damit eine Abdeckung sichergestellt ist.

Tabelle 1: Identifiziertes Set potenzieller Confounder inklusive vorgeschlagener Operationalisierungen in der Indikation RRMS (mehreseitige Tabelle)

Nr.	potenzieller Confounder ^a	vorgeschlagene Operationalisierung
1	Geschlecht	weiblich vs. männlich vs. divers
2	Alter	stetig
3	EDSS Score	stetig
4	EDSS Veränderung im Vorjahr	stetig
5	Multiple Sclerosis Functional Composite (MSFC) Score	stetig
6	Fatigue	mithilfe eines validierten Instruments: stetig
7	Komorbiditäten	Anzahl der Komorbiditäten ^b : stetig
8	Absolute Lymphozytenzahl; Lymphopenie	absolute Lymphozytenzahl: stetig
9	Studienzentrum; Zentrum für Multiple Sklerose	MS-spezialisiertes vs. nicht spezialisiertes Zentrum ^b
10	BMI	stetig
11	Tabakkonsum	ja vs. nein <u>oder</u> Raucher vs. ehemaliger Raucher vs. Nieraucher
12	infratentorielle Läsionen	ja vs. nein
13	Anzahl der Kontrastmittel-anreichernden Läsionen	stetig
14	Volumen T1-hypointensiver Läsionen	stetig
15	Anzahl der T2 Läsionen	stetig
16	Volumen der T2 Läsionen	stetig
17	Verträglichkeit der Vortherapie und Erwartungen an die Behandlung ^c	keine adäquate Operationalisierung möglich, siehe Ausführungen im folgenden Textabschnitt
18	vorherige Behandlung mit einer anderen krankheitsmodifizierenden Therapie	ja vs. nein
19	Patientenwunsch	keine adäquate Operationalisierung möglich, siehe Ausführungen im folgenden Textabschnitt
20	gesundheitsbezogene Lebensqualität	mithilfe eines validierten Instruments: stetig
21	Anzahl der vorherigen Schübe	Schübe im letzten Jahr: stetig
22	Schwere der Schübe	mit Steroiden behandelter Schub im letzten Jahr: ja vs. nein
23	Zeit seit der Diagnose der Multiplen Sklerose	Zeit seit der MS-Diagnose: stetig
24	Cerebrospinalflüssigkeit (CSF) Status	Liquor spezifische oligoklonale Banden: ja vs. nein
25	Art der beruflichen Tätigkeit	körperliche Arbeit: ja vs. nein

Tabelle 1: Identifiziertes Set potenzieller Confounder inklusive vorgeschlagener Operationalisierungen in der Indikation RRMS (mehrseitige Tabelle)

Nr.	potenzieller Confounder ^a	vorgeschlagene Operationalisierung
26	Bildungsgrad	Bildungszeit in Jahren: stetig <u>oder</u> kein Studienabschluss oder keine höhere Ausbildung vs. Studienabschluss oder höhere Ausbildung
27	Abstammung	weiß, asiatisch, schwarz oder afroamerikanisch, andere
28	Kürzliche Schwangerschaft (in den letzten 3 Monaten)	ja vs. nein
<p>a. Die durch Zusammenführung reduzierten Confounder wurden frei von Englisch auf Deutsch übersetzt. Alle identifizierten potenziellen Confounder-Variablen inklusive der Zuordnung zu den jeweiligen potenziellen Confoundern sind in Anhang B, Tabelle 2 und Tabelle 3 zu entnehmen. Dort wurde auf die Übersetzung der Confounder-Variablen verzichtet, um Übersetzungsfehler zu vermeiden.</p> <p>b. Diese Operationalisierung wurde unter anderem im Workshop mit den externen Sachverständigen diskutiert, weitere detaillierte Informationen sind dem nachfolgenden Abschnitt zur Operationalisierung der potenziellen Confounder zu entnehmen.</p> <p>c. Diese Variable beinhaltet die Komponenten: Gründe für den Abbruch der vorherigen Therapie, Wirksamkeit (bekannt oder erwartet), Verträglichkeit, Nebenwirkungen anderer Medikationen</p> <p>BMI: Body-Mass-Index; EDSS: Expanded Disability Status Scale; MS: multiple Sklerose</p>		

Von diesen 28 potenziellen Confoundern konnten 26 mithilfe der Literatur und 18 über die Interviews mit klinischen Expertinnen und Experten identifiziert werden. Die potenziellen Confounder Patientenwunsch und Art der beruflichen Tätigkeit wurden nicht anhand der Literatur, sondern nur aufgrund der Interviews identifiziert.

Operationalisierung der potenziellen Confounder

Zuletzt wurden geeignete Operationalisierungen der potenziellen Confounder vorgeschlagen, welche ebenfalls in Tabelle 1 dargestellt sind. Falls möglich, wurde eine stetige Operationalisierung der Confounder bevorzugt, da diese den höchsten Informationsgehalt bietet. Zum Beispiel wurde die stetige Operationalisierung des EDSS-Scores der dichotomen Operationalisierung basierend auf einem Trennwert von 3,5 ($\leq 3,5$ vs. $> 3,5$) vorgezogen. Darüber hinaus kann geprüft werden, ob ordinalskalierte Confounder als stetige Variable aufgefasst werden können.

Sind stetige Operationalisierungen nicht möglich, wird eine kategorielle Operationalisierung gewählt. Hierbei ist es aufgrund des höheren Informationsgehalts zu bevorzugen, dass die Ausprägungen nicht zusammengefasst werden. Nur für den Fall, dass statistische Modelle nicht konvergieren, können Zusammenfassungen von Ausprägungen (z. B. Dichotomisierungen) in Betracht gezogen werden. Diese Zusammenfassungen sind zu präspezifizieren. Zum Beispiel könnte eine Dichotomisierung des Confounders Tabakkonsum (ja vs. nein) präspezifiziert werden, falls in einer Kategorie (Raucher vs. ehemaliger Raucher

vs. Nieraucher) zu wenige Patientinnen und Patienten eingehen und das statistische Modell nicht konvergiert.

Für die beiden Confounder Patientenwunsch sowie Verträglichkeit der Vortherapie und Erwartungen an die Behandlung war keine adäquate Operationalisierung möglich (siehe hierzu die Ausführungen im folgenden Abschnitt).

Ergebnisse des Workshops mit den externen Sachverständigen

Die Confounder-Liste mit den ausgewählten Operationalisierungen wurde in einem Workshop mit den externen Sachverständigen diskutiert. Im Folgenden werden übergeordnet die im Workshop diskutierten Punkte und die daraus gezogenen Schlussfolgerungen zu den in Tabelle 1 dargestellten Confoundern wiedergegeben.

Nicht messbare Confounder

Ein wesentlicher Aspekt, der im Workshop diskutiert wurde, war die Operationalisierung des Patientenwunschs. Dieser ist komplex und beinhaltet die persönlichen Erwartungen sowie bereits gemachte individuelle Erfahrungen der Patientin bzw. des Patienten. Laut den externen Sachverständigen sind Interaktionen zwischen den verschiedenen Parametern, die den Patientenwunsch bestimmen, zu erwarten und der Patientenwunsch sollte möglichst granular operationalisiert werden. Den Patientenwunsch im Sinne einer partizipativen Therapieentscheidung („shared decision making“) mit all den unterschiedlichen, individuellen Gesichtspunkten adäquat abzubilden, scheint in dieser Form nicht möglich zu sein. Dies trifft teilweise genauso auf die Faktoren zu, die die Ärztin bzw. den Arzt in ihrer Therapieentscheidung beeinflussen. Beispielsweise kann die Arztentscheidung durch Gründe für einen vorherigen Therapieabbruch, vermutete oder bekannte Arzneimittelwirkungen, die Verträglichkeit des Arzneimittels oder Nebenwirkungen anderer Medikamente gelenkt werden. Der potenzielle Confounder Verträglichkeit der Vortherapie und Erwartungen an die Behandlung spiegelt beispielsweise einen solchen Faktor wider. Analog zu dem Patientenwunsch ist auch hier aufgrund der Komplexität und Vielschichtigkeit eine adäquate Operationalisierung nicht möglich. Die Anregung der externen Sachverständigen, ggf. eher die Beweggründe für eine nächste Therapie aufzuschlüsseln, kann aufgrund der Vielzahl hinzukommender, verschiedenster Variablen im Sinne der Machbarkeit nicht berücksichtigt werden.

Da eine Operationalisierung und damit einhergehend eine Erhebung der beiden identifizierten potenziellen Confounder Patientenwunsch sowie Verträglichkeit der Vortherapie und Erwartungen an die Behandlung nicht möglich ist, können die beiden als relevant eingestuften potenziellen Confounder nicht in die Auswertung einbezogen werden (nicht messbare Confounder, siehe Glossar). Für diese potenziellen Confounder kann in der nicht

randomisierten vergleichenden Studien nicht adjustiert werden. Dies muss anschließend bei der Ergebnisinterpretation berücksichtigt werden [10].

Potenzielle Confounder, welche nicht messbar oder nicht operationalisierbar sind, sollten stets benannt werden, und es sollte klar und nachvollziehbar dargelegt werden, welche Variablen hierunter fallen und warum eine Operationalisierung nicht möglich ist (siehe auch Abschnitt 5.3 zu den Handlungsempfehlungen). Zudem muss diskutiert werden, was die Nichtberücksichtigung dieser Confounder für die Interpretation der Ergebnisse bedeutet.

Komorbiditäten

Im Workshop mit den externen Sachverständigen wurde des Weiteren über den Confounder Komorbiditäten diskutiert. Anstelle eines einzelnen zusammenfassenden Confounders Komorbiditäten würden die externen Sachverständigen befürworten, wichtige Komorbiditäten (wie z. B. Psoriasis Arthritis oder gastrointestinale Erkrankungen) als separate Confounder aufzunehmen. Durch die Sammelkategorie würden Informationen verloren gehen. Beispielsweise würde bei Vorliegen einer Psoriasis Arthritis eher Dimethylfumarat und bei gastrointestinalen Beschwerden oder einer Lymphopenie eher Glatirameracetat verordnet. Diese Empfehlung der externen Sachverständigen ist grundsätzlich nachvollziehbar und sollte nach Möglichkeit auch umgesetzt werden, um möglichst wenige Informationen zu verlieren. Denn grundsätzlich ist eine Zusammenfassung mit der Annahme verbunden, dass alle berücksichtigten Komorbiditäten die gleiche Relevanz haben hinsichtlich Stärke und Richtung der Verzerrung bei Nichtberücksichtigung. Komorbiditäten einzeln zu betrachten ist insbesondere dann sinnvoll, wenn nur wenige Komorbiditäten für eine Fragestellung von Bedeutung sind. Dennoch ist die Empfehlung vor dem Hintergrund der Anzahl ($n = 18$) an identifizierten Komorbiditäten im vorliegenden Fall zu diskutieren. Je nach für die statistischen Analysen zur Verfügung stehender Patientenzahl ist abzuwägen, ob der durch die Zusammenführung in Kauf genommene Informationsverlust vertretbar ist. Im vorliegenden Fall wird die Zusammenfassung der verschiedenen Komorbiditäten zu einem Confounder Anzahl an Komorbiditäten in der Gesamtschau als vertretbar angesehen.

Studienzentrum

Zudem wurde zu der gewählten Operationalisierung des potenziellen Confounders Studienzentrum über spezialisierte MS-Zentren vs. nicht spezialisierte MS-Zentren von den externen Sachverständigen geäußert, dass die Studienzentren im Sinne eines Stratifizierungsmerkmals betrachtet werden sollten. Die externen Sachverständigen sehen jedes einzelne Studienzentrum als relevanten Confounder an, da abhängig vom jeweiligen Zentrum stark unterschiedliche Kapazitäten z. B. für Infusionen oder Magnetresonanztomografien (MRTs) zur Verfügung ständen und teilweise unterschiedliche Behandlungsphilosophien gelebt werden würden. Dies könne einen potenziellen Einfluss auf das Behandlungsergebnis haben. Diese Empfehlung scheint in Anbetracht der Anzahl an deutschen Zentren, an denen

Patientinnen und Patienten mit multipler Sklerose behandelt werden, und der dadurch stark ansteigenden Anzahl im statistischen Modell zu berücksichtigender Ausprägungen nicht umsetzbar. In Abhängigkeit von der zur Verfügung stehenden Patientenzahl ist zu diskutieren, ob es neben der in Tabelle 1 vorgeschlagenen Operationalisierung ggf. weitere geeignete Operationalisierungen gibt.

Zusammenfassung

Insgesamt konnten die von den externen Sachverständigen identifizierten 160 potenziellen Confounder-Variablen auf 28 potenzielle Confounder zusammengeführt werden. Hierbei lag der Fokus insbesondere auf der methodischen Vorgehensweise und der notwendigen transparenten Darstellung dieses Vorgehens. Die potenziellen Confounder-Variablen wurden mit dem Ziel, Duplikate und Überlappungen zu reduzieren, zusammengefasst. Damit sollte höchstens ein vertretbarer Informationsverlust hinsichtlich der identifizierten potenziellen Confounder einhergehen. Aus den gewonnenen Erfahrungen werden übergeordnete Handlungsempfehlungen abgeleitet, welche in Abschnitt 5.3 beschrieben werden.

Ähnlich wie in Pufulete 2022 beschrieben, war eine Auswahl der Confounder auf Basis eines kausalen Verständnisses der Zusammenhänge der verschiedenen Variablen nicht das Ziel der hier beschriebenen Confounderidentifikation und -zusammenführung. Inhaltlich überlappende Confounder-Variablen, die wahrscheinlich zueinander hoch korreliert sind, sollten zusammengeführt worden sein. Ob es sich bei den identifizierten potenziellen Confoundern ggf. um Mediatoren oder Collider handelt (siehe Glossar), wurde nicht überprüft. Der Ausschluss solcher Variablen sollte literaturgestützt begründet werden und die wechselseitigen Beziehungen der potenziellen Confounder mithilfe eines gerichteten azyklischen Graphen (DAG) [1,2] visualisiert werden.

5.2.2 Weitere Vorschläge zur systematischen Confounderidentifikation in der Literatur

Abbildung 3 zeigt das Ergebnis der fokussierten Informationsbeschaffung nach relevanten Publikationen zur systematischen Confounderidentifikation gemäß den Kriterien für den Studieneinschluss (siehe Abschnitt 4.2.2). Die Suchstrategie für die Suche in MEDLINE findet sich in Abschnitt A.2. Die Suche fand am 13.12.2024 statt.

Die Referenzen der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich in Abschnitt A.1.

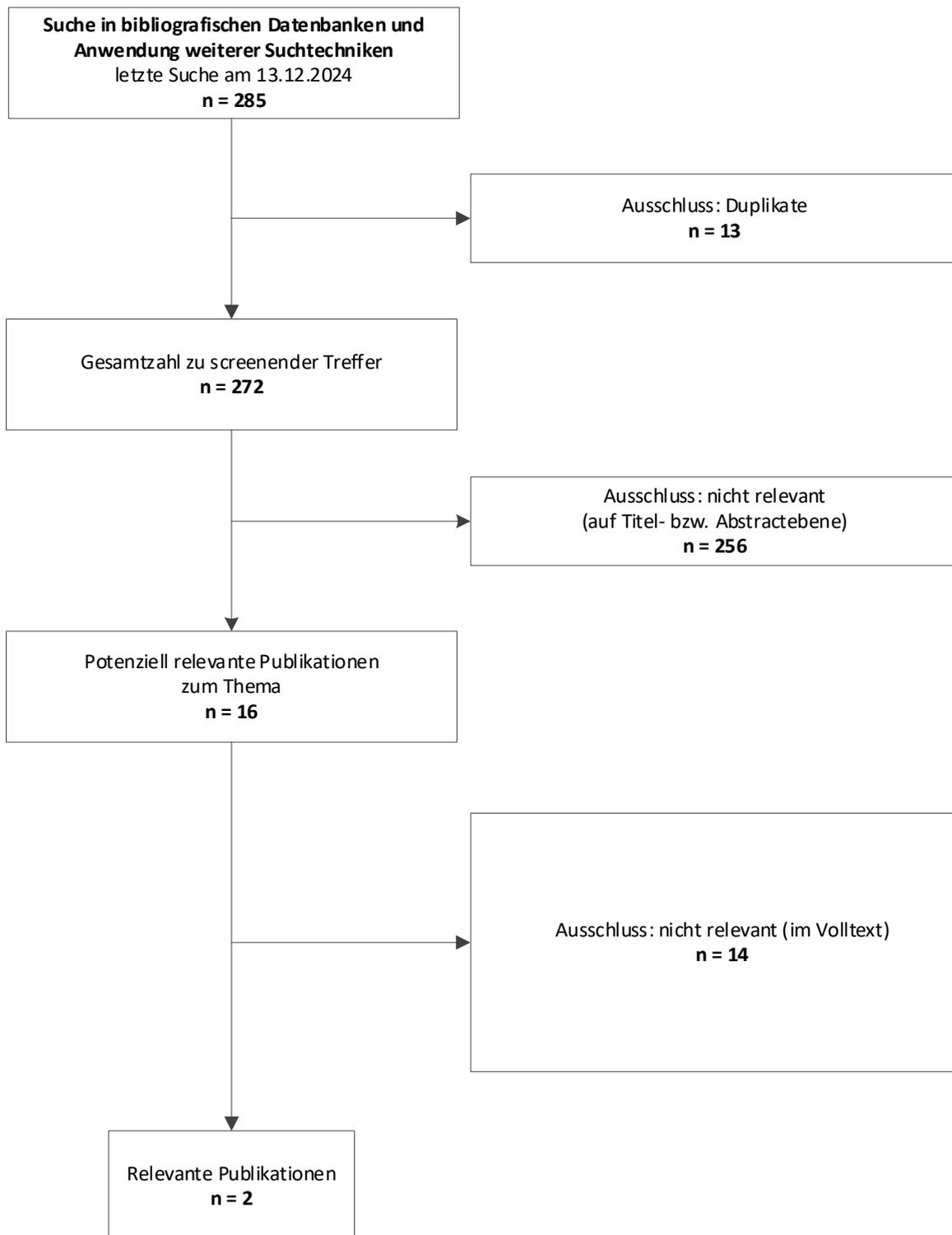


Abbildung 3: Ergebnis der fokussierten Informationsbeschaffung und der Studienselektion nach Publikationen zur systematischen Vorgehensweise der Confounderidentifikation

Die 2 als relevant eingestuft Publications umfassen die Publikation von Pufulete 2022 [11] selbst und einen HTA-Report von Harris und Pufulete 2023 [27], der die von Pufulete

beschriebene Methodik der Confounderidentifikation beinhaltet. Darüberhinausgehende Publikationen zu systematischen Ansätzen der Confounderidentifikation wurden nicht identifiziert.

5.3 Vorschlag für die Anforderung an eine systematische Confounderidentifikation

Zur Evaluation von Interventionseffekten auf Basis eines nicht randomisierten Vergleichs, sollte im Rahmen der Studienplanung a priori ein Set potenzieller Confoundern identifiziert werden. Hierfür wurden basierend auf den Ergebnissen einer Confounderidentifikation durch externe Sachverständige und anhand der weitergeführten eigenen Untersuchungen konkrete Handlungsempfehlungen zur systematischen Identifikation von Confoundern abgeleitet. Diese stellen einen Vorschlag für die konkreten Anforderungen an die systematische Confounderidentifikation dar.

Auch wenn sich der Bericht der externen Sachverständigen (ebenso wie auch die Arbeit von Pufulete 2022) auf eine Indikation mit umfangreicher Studienlage bezieht, wurde bei der Ableitung der Handlungsempfehlungen auch die Situation der selten(er)en Erkrankungen berücksichtigt, da dies die regelhafte Situation z. B. im Rahmen der anwendungsbegleitenden Datenerhebung darstellt. Sollten sich hierfür besondere Anforderungen ergeben, wird dies im Folgenden explizit beschrieben.

Übergeordnet gliedert sich der erarbeitete Vorschlag zur systematischen Confounderidentifikation in 3 Schritte, die in Abbildung 4 dargestellt sind.

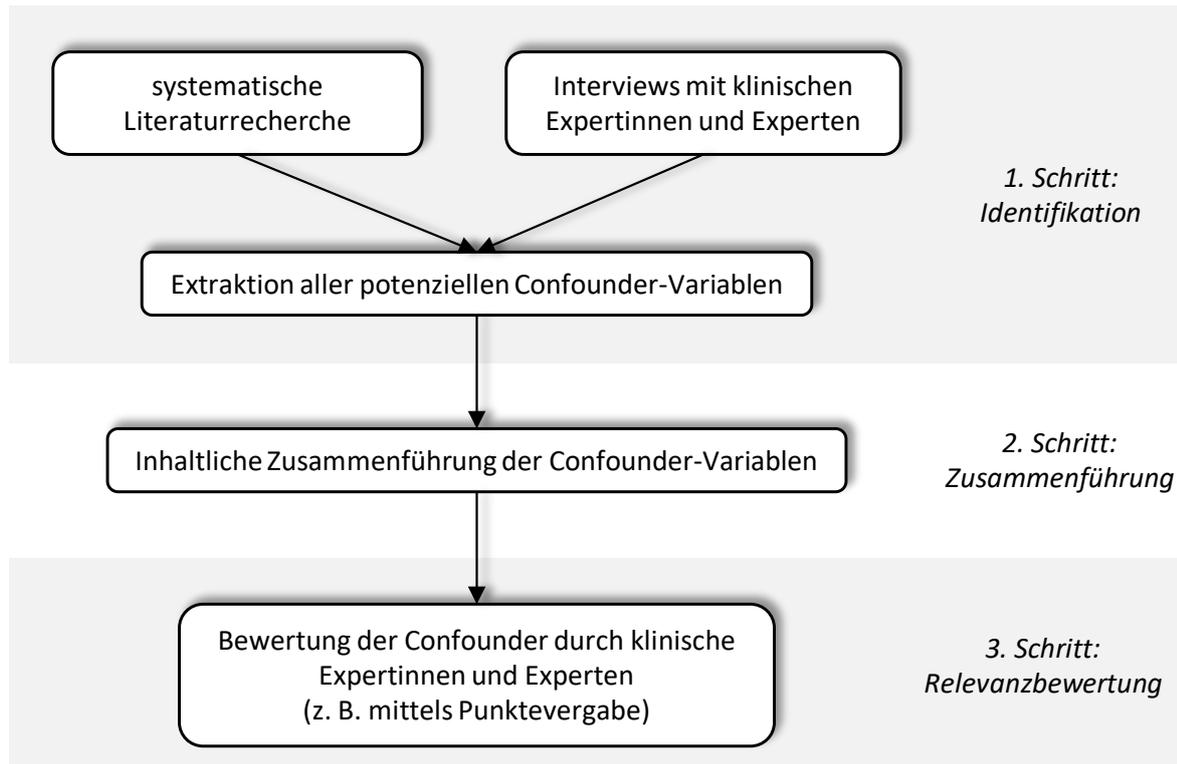


Abbildung 4: Vorgeschlagenes 3-schrittiges Verfahren der systematischen Confounderidentifikation

Angelehnt an Pufulete 2022 setzt sich der 1. Schritt der Identifikation aus 2 Bausteinen zusammen: einer systematischen Literaturrecherche und Interviews mit klinischen Expertinnen und Experten. Nach der erfolgten Extraktion aller potenziellen Confounder-Variablen wird im 2. Schritt im Sinne der Aufwandsreduktion eine Zusammenführung inhaltlich überlappender Confounder-Variablen vorgenommen. Dieser Schritt stellt eine Weiterentwicklung im Vergleich zu der in Pufulete 2022 beschriebenen Methodik dar. Anschließend erfolgt im 3. Schritt eine Relevanzbewertung der bereits zusammengeführten, potenziellen Confounder hinsichtlich ihrer Relevanz durch die klinischen Expertinnen und Experten.

Um die Confounderidentifikation nachvollziehen zu können, ist eine detaillierte und transparente Dokumentation der Vorgehensweise und der Ergebnisse erforderlich. Diese Dokumentation sollte als Teil des Studienprotokolls der geplanten nicht randomisierten vergleichenden Studie veröffentlicht werden.

Nachfolgend werden für die einzelnen Schritte des Verfahrens die identifizierten Handlungsempfehlungen erläutert. Die einzelnen Empfehlungen sind auch als Checkliste in Anhang C zu finden.

5.3.1 Schritt 1: Identifikation

Bei der systematischen Informationsbeschaffung sollte auf Primärpublikationen zu RCTs und Kohortenstudien zurückgegriffen werden. Nachfolgend können auch Leitlinien oder 1-armige Studien berücksichtigt werden.

In Abhängigkeit von der Qualität und der Anzahl der vorliegenden Publikationen in dem zu untersuchenden Anwendungsgebiet sollte über den einzuschließenden Studientyp entschieden werden. Zunächst sollten analog zu Pufulete 2022 und den externen Sachverständigen geeignete RCTs und Kohortenstudien einbezogen werden. Insbesondere für Fragestellungen in Indikationen mit absehbar weniger umfassender Studienlage (z. B. bei seltenen Erkrankungen), kann beim Fehlen von geeigneten RCTs und Kohortenstudien auf 1-armige Studien (inklusive Vorher-Nachher-Vergleichen) und die darin in tabellarischer Form aufgeführten Patientencharakteristika zurückgegriffen werden. Genauso besteht die Möglichkeit in diesen Fällen auch z. B. Leitlinien in die Recherche einzubeziehen und diese nach Prognose- oder Risikofaktoren ebenso wie Faktoren, die die Therapieentscheidung bestimmen, zu durchsuchen.

Ein alleiniger Rückgriff auf Sekundärliteratur wird nicht empfohlen. Dieser wäre nur unter der Voraussetzung möglich, dass die Sekundärliteratur hochwertig und aktuell ist, die komplette Fragestellung abdeckt, geeignete Primärstudien in ausreichendem Umfang eingehen und die für die Bewertung der Confounder notwendigen Informationen (bspw. Adjustierungsvariablen) in ausreichendem Umfang berichtet werden. Die bisherigen Erfahrungen aus der Prüfung der Studienunterlagen im Rahmen der anwendungsbegleitenden Datenerhebung zeigen, dass dies in der Regel nicht der Fall ist [21,22].

Die Ein- und Ausschlusskriterien für den Studienpool zur Confounderextraktion sollten sich nah an der zu untersuchenden Fragestellung orientieren.

Die Ein- und Ausschlusskriterien der Literaturrecherche sollten auf die konkrete Fragestellung zugeschnitten sein. Es empfiehlt sich, die Ein- und Ausschlusskriterien anhand der Population, Intervention, Comparison, Outcome (PICO) der zu untersuchenden Fragestellung zu definieren. Insbesondere die Definition der Population (P) sollte dabei nah an der zu untersuchenden Fragestellung liegen. Zum Beispiel liegen für die vorliegende Untersuchung Publikationen, die Schübe während der Schwangerschaft untersuchen, außerhalb der Fragestellung der nachzubildenden Studie CONFIRM und müssen entsprechend für die Confounderextraktion nicht berücksichtigt werden.

Hinsichtlich der Intervention und des Komparators muss in der Regel für den Studieneinschluss nur eine der beiden Komponenten erfüllt sein. So können nachvollziehbarerweise auch Studien eingeschlossen werden, die ausschließlich den Komparator (das „C“ des PICO) nicht aber die Intervention („I“ des PICO) untersuchen.

Einen konkreten Endpunkt (O) in die Einschlusskriterien aufzunehmen, empfiehlt sich vor allem dann, wenn die zu untersuchende Fragestellung explizit auf diesen Endpunkt abstellt. Für Fragestellungen der anwendungsbegleitenden Datenerhebung bzw. der Nutzenbewertung von Arzneimitteln ist eine solche Einschränkung in der Regel nicht notwendig, da eine Vielzahl von patientenrelevanten Endpunkten Gegenstand der Untersuchung ist.

Die Extraktion sollte bei vergleichenden Beobachtungsstudien alle Adjustierungsvariablen – unabhängig von der statistischen Signifikanz – umfassen. Bei randomisierten oder nicht vergleichenden Studien sollten alle Patientencharakteristika extrahiert werden.

Analog zum Vorgehen von Pufulete 2022 und den externen Sachverständigen sollten als potenzielle Confounder alle Faktoren, die zur Adjustierung in den statistischen Analysen verwendet wurden, extrahiert werden. Bei RCTs oder je nach Datenlage 1-armigen Studien sollten die in tabellarischer Form berichteten Patientencharakteristika („Tabelle 1“) extrahiert werden. Hierbei ist es insbesondere von Bedeutung, dass transparent und nachvollziehbar ist, welche potenzielle Confounder-Variable aus welcher Quelle (Literaturangabe, idealerweise inkl. Angabe des konkreten Zitats) extrahiert wurde, z. B. in einem Anhangsdokument zum Hauptteil der Confounderidentifikation.

Sollte in einer Publikation ein DAG identifiziert werden, so sollten die im DAG präsentierten Variablen extrahiert werden. Dies ist im erwartbar höheren Informationsgehalt des DAG begründet. So werden im DAG auch solche Variablen enthalten sein, die im statistischen Modell nicht zur Adjustierung aufgenommen wurden, da diese im Datensatz nicht erhoben sind.

Nach der klassischen Definition eines Confounders beeinflusst dieser sowohl die Behandlung als auch den zu untersuchenden Endpunkt [1,2]. Ausgehend von dieser Definition sollte für jeden zu untersuchenden Endpunkt (im Bericht der externen Sachverständigen die jährliche Schubrate, die bestätigte Behinderungsprogression anhand EDSS, gesundheitsbezogene Lebensqualität und unerwünschte Ereignisse [UEs]) ein eigenes Set an potenziell relevanten Confoundern identifiziert werden. Insbesondere für Auswertungen, in denen die Confounder-Adjustierung mittels Propensity Score Verfahren (wie Matching oder Gewichtung) geplant sind, ist es aus pragmatischen Gründen ausreichend, eine übergeordnete Menge aller potenziell relevanten Confounder über alle Endpunkte, zu definieren, welche für die Auswertung aller Endpunkte eingesetzt werden kann.

Bei der Extraktion aus Primärpublikationen wird die Anwendung eines geeigneten Sättigungskriteriums empfohlen.

In Anlehnung an das Vorgehen in der Publikation von Pufulete 2022 und den externen Sachverständigen ist es sinnvoll, vorab ein objektives und plausibles Sättigungskriterium für die Extraktion der Primärpublikationen zu definieren. Pufulete 2022 verwendete ein

Sättigungskriterium, dass die Extraktion gestoppt wird, wenn innerhalb von 10 konsekutiven Studien kein neuer potenzieller Confounder identifiziert wurde. Die externen Sachverständigen wendeten nur für Studien vor 2020 ein Sättigungskriterium an und stoppten die Extraktion aus den zufällig ausgewählten Studien, wenn in 15 konsekutiven Studien kein neuer Confounder gefunden wurde. Die Verwendung eines Sättigungskriteriums ist insbesondere bei einer hohen Anzahl identifizierter Publikationen sinnvoll, um Ressourcen einzusparen. Es wird davon ausgegangen, dass die Verwendung von 10 konsekutiven Studien als Sättigungskriterium angemessen ist, da zur Identifikation potenzieller Confounder auch weitere Quellen, bspw. qualitative Interviews mit klinischen Expertinnen und Experten verwendet werden.

Es sollen Interviews mit klinischen Expertinnen und Experten durchgeführt werden.

Die Identifikation potenzieller Confounder durch klinische Expertinnen und Experten stellt einen wichtigen Baustein des vorgeschlagenen 3-schrittigen Verfahrens dar.

Der Ablauf der Interviews sollte transparent dokumentiert sein. So entwickelten Pufulete 2022 sowie die externen Sachverständigen z. B. Vignetten-basierte Interviewleitfäden, in denen verschiedene Szenarien zu klinischen Fällen beschrieben wurden (siehe Supplement zu Pufulete 2022 [11] oder Anhang B des Berichts der externen Sachverständigen). Es ist hervorzuheben, dass in diesem Schritt nicht nur eine Validierung der bereits identifizierten potenziellen Confounder aus der Literaturrecherche stattfinden sollte, sondern die Einbindung der Identifikation weiterer potenzieller Confounder dient.

Für die Zahl der einzubindenden klinischen Expertinnen und Experten kann keine eindeutige Empfehlung gegeben werden. Pufulete 2022 haben je Indikation Interviews mit 6 klinischen Expertinnen und Experten (also insgesamt 12) geführt und gehen davon aus, dass dies ausreichend ist, um eine Sättigung zu erreichen. Die externen Sachverständigen haben 8 klinische Expertinnen und Experten eingebunden. Wie viele Interviews tatsächlich nötig sind, um eine (theoretische) Sättigung zu erreichen, sollte indikationsabhängig eingeschätzt werden. Als Richtwert kann hierbei eine Anzahl von 9 bis 17 Interviews dienen, mit der im Rahmen der qualitativen Forschung fragestellungsübergreifend eine Sättigung erreicht werden konnte [33]. Dies deckt sich mit der Herangehensweise von Pufulete 2022 und den externen Sachverständigen. Entscheidungskriterien für die Anzahl an durchzuführenden Interviews sind z. B. die Indikation, die erwartbare Studienlage oder die Variabilität im Erkrankungsverlauf.

In der vorliegenden Indikation RRMS wurden 35/160 (22 %) der potenziellen Confounder-Variablen durch die Einbindung von klinischen Expertinnen und Experten identifiziert, welche nicht in der Literatur gefunden wurden. In der zusammengeführten Liste an potenziellen Confoundern verblieben noch 2/28 (7 %) Confounder, die ausschließlich in den Interviews

benannt wurden. In der Publikation von Pufulete 2022 wurden 11/70 (16 %) der potenziellen Confounder-Variablen nicht über die Literatursuche identifiziert, von diesen 7 (10 %) ausschließlich über qualitative Interviews.

Auch wenn der Anteil über Interviews identifizierter Confounder in den vorliegenden Beispielen eher gering erscheint, sollte nicht die Schlussfolgerung getroffen werden, dass dieser Schritt verzichtbar ist. Zum einen ist absehbar, dass in Indikationen, in denen ggf. eine geringere Menge an Publikationen zur Verfügung steht, den Interviews eine größere Bedeutung zuteilwird. Zum anderen haben sowohl Pufulete 2022 als auch die externen Sachverständigen über die Interviews insbesondere patientenzentrierte Faktoren identifiziert, die in der Literatur bisher nicht berücksichtigt werden. Auch wenn es sich hierbei um schwer oder gar nicht zu operationalisierende Variablen handelt, trägt das Wissen um die Bedeutung dieser Faktoren zu einer differenzierteren Beschreibung der Black-Box des „unmeasured confounding“ (siehe Glossar) bei. Dies sollte im Anschluss bei der Interpretation der Ergebnisse des adjustierten nicht randomisierten Vergleichs berücksichtigt werden.

Die klinischen Expertinnen und Experten sollten namentlich benannt und potenzielle Interessenkonflikte offengelegt werden.

Es sollte transparent dokumentiert werden, wie viele klinische Expertinnen und Experten mit welchem Fachgebiet eingebunden wurden. Für die Dokumentation ist es erforderlich die klinischen Expertinnen und Experten namentlich zu benennen und potenzielle Interessenkonflikte offenzulegen (siehe Anhang C [Handlungsempfehlungen Checkliste]). Die Dokumentation kann z. B. in einem Anhangsdokument zum Hauptteil der Confounderidentifikation vorgelegt werden. Die Kenntnis des Studienprotokolls der geplanten Studie, für die die Confounderidentifikation durchgeführt wird oder auch die Kenntnis des Datensatzes der zur Verfügung stehenden Datenquelle kann die Einschätzung der klinischen Expertinnen und Experten beeinflussen. Im Fall der anwendungsbegleitenden Datenerhebung sollten die klinischen Expertinnen und Experten (oder zumindest ein Teil derer) deshalb idealerweise keine Kenntnis des Datensatzes der zur Verfügung stehenden Datenquelle (z. B. des Indikationsregisters) haben.

5.3.2 Schritt 2: Zusammenführung

Inhaltlich überlappende Confounder-Variablen sollten zusammengeführt werden. Dies sollte nach der Extraktion erfolgen. Eine transparente und nachvollziehbare Dokumentation sollte gewährleistet sein.

Nach der Extraktion aller identifizierten potenziellen Confounder-Variablen aus 1) der systematischen Literaturrecherche und 2) den qualitativen Interviews mit den klinischen Expertinnen und Experten wird empfohlen – wo sinnvoll – eine Zusammenführung von duplizierten bzw. stark überlappenden potenziellen Confounder-Variablen vorzunehmen. Hierbei werden potenzielle Confounder-Variablen, welche ein inhaltlich gleiches Konstrukt

messen oder sich inhaltlich überlappen, zusammengefasst. Dabei ist eine entsprechende Expertise im Anwendungsgebiet zu berücksichtigen. Dieses Vorgehen führt dazu, dass potenzielle Confounder-Variablen, welche stark miteinander korrelieren, gebündelt betrachtet werden.

Bei der Zusammenführung können identifizierte potenzielle Confounder-Variablen unter verschiedene potenzielle Confounder fallen. Eine Zuordnung zu einem der potenziellen Confounder reicht aus, da dadurch die Abdeckung sichergestellt ist.

Die Zusammenführung sollte nach der Extraktion aller identifizierten potenziellen Confounder-Variablen erfolgen, um eine transparente und nachvollziehbare Dokumentation zu gewährleisten. Es muss nachvollziehbar sein, welche potenziellen Confounder-Variablen zu welchen potenziellen Confoundern zusammengeführt wurden.

Den Schritt der Zusammenführung überlappender potenzieller Confounder-Variablen bereits vor der anschließenden Relevanzbewertung der potenziellen Confounder durch klinische Expertinnen und Experten durchzuführen ist sinnvoll, um den Bewertungsaufwand eines jedes einzelnen Confounders für die Klinikerinnen und Kliniker möglichst gering zu halten ohne dabei einen relevanten Informationsverlust in Kauf nehmen zu müssen.

Für jede identifizierte Confounder-Variable sollte geprüft werden, ob sie innerhalb der zu untersuchenden Fragestellung liegt.

Aufgrund der intensiven inhaltlichen Auseinandersetzung mit jeder einzelnen identifizierten potenziellen Confounder-Variable bietet es sich an, im Rahmen des Schrittes der Zusammenführung, für jede Variable zu prüfen, ob sie innerhalb der zu untersuchenden Fragestellung liegt und die methodischen und inhaltlichen Kriterien eines Confounders für die zu untersuchende Fragestellung erfüllt:

- Bei jeder Confounderidentifikation, welche bspw. im Rahmen einer anwendungsbegleitenden Datenerhebung durchgeführt wird, liegt eine zuvor definierte, zu untersuchende Fragestellung vor, für die analog zu einer RCT Ein- und Ausschlusskriterien festgelegt werden. Ist eine Variable in der vorliegenden Fragestellung ein Ausschlusskriterium, muss diese im weiteren Verlauf nicht weiter berücksichtigt werden.
- Im Sinne der Nachbildung einer randomisierten Studie dürfen nur Variablen, die bereits zu Studienbeginn (Baseline) bekannt sind, als Confounder berücksichtigt werden. Jegliche Variablen, die Studienverlaufscharakteristika darstellen, müssen ausgeschlossen werden.
- Variablen, die in der zu untersuchenden Fragestellung lediglich 1 Ausprägung annehmen können (z. B. müssen alle Patientinnen und Patienten anhand eines konkreten

Diagnosekriteriums identifiziert worden sein) stellen für diese Fragestellung keinen Confounder dar. Um diese Variablen zu identifizieren, bietet sich der Abgleich mit den konkreten Ein- und Ausschlusskriterien der geplanten Studie an.

- Soll die Fragestellung anhand von versorgungsnahen Daten aus deutschen Studienzentren untersucht werden, können Variablen, wie „vorherige Behandlungskosten“ basierend auf der fehlenden Relevanz im deutschen Versorgungskontext ausgeschlossen werden.
- Ist eine prospektive Datenerhebung geplant, können alle Variablen ausgeschlossen werden, die nur für eine retrospektive Datenerhebung relevant sind (z. B. Zeit seit der letzten Visite).

Der Ausschluss eines potenziellen Confounders aus inhaltlichen Überlegungen muss literaturgestützt begründet werden.

Die Annahmen, die zur Identifikation einer Variable als potenzieller Confounder getroffen werden, sind in der Regel mit Unsicherheiten behaftet. Ebenso wird das Wissen um die Zusammenhänge der verschiedenen identifizierten Variablen untereinander nie vollständig sein. Daher ist es notwendig, für den Ausschluss einer Variable aus der Liste der identifizierten potenziellen Confounder-Variablen eine literaturgestützte Begründung vorzulegen. Die Einstufung einer Variable von den eingebundenen klinischen Expertinnen und Experten als unwichtig ist nicht ausreichend. Ebenso ist das Fehlen eines geeigneten Instruments zur Erhebung eines potenziellen Confounders keine Begründung für den Ausschluss eines potenziellen Confounders (siehe auch Ausführungen unten zur Operationalisierung).

Grundsätzlich sollte im statistischen Modell eine Adjustierung für potenzielle Mediatoren oder Collider vermieden werden (siehe z. B. Williams 2018 [2]). Ein Ausschluss von Variablen sollte grundsätzlich literaturgestützt erfolgen. Liegt lediglich der Verdacht vor, dass es sich bei einer Variablen um bspw. einen Collider handelt, kann diese im Rahmen einer Sensitivitätsanalyse aus dem Modell ausgeschlossen werden.

Confounder, deren Relevanz unklar ist, sollten erhoben und in der statistischen Analyse berücksichtigt werden.

Für alle identifizierten potenziellen Confounder ist eine Operationalisierung festzulegen.

Für alle identifizierten potenziellen Confounder ist im Studienprotokoll bzw. statistischen Analyseplan a priori eine geeignete Operationalisierung festzulegen.

Basierend auf dem Informationsgehalt sind stetige Operationalisierungen zu bevorzugen. Sind stetige Operationalisierungen nicht möglich, wird eine kategorielle Operationalisierung gewählt. Hierbei ist es aufgrund des höheren Informationsgehalts zu bevorzugen, dass die Ausprägungen nicht zusammengefasst werden. Nur für den Fall, dass das statistische Modell

nicht konvergiert, können Zusammenfassungen von Ausprägungen (z. B. Dichotomisierungen) in Betracht gezogen werden. Diese Zusammenfassungen sind zu präzisieren.

Wurden Variablen als potenzielle Confounder identifiziert, die anhand eines Instruments für patientenberichtete Endpunkte gemessen werden, wie z. B. die gesundheitsbezogene Lebensqualität oder Fatigue, sollte die Operationalisierung sicherstellen, dass diese mithilfe eines validierten Instruments erhoben werden.

Prinzipiell sollte auch die geeignete Operationalisierung im Sinne der Machbarkeit gewählt werden. Zum Beispiel ist es nachvollziehbar, wenn die Operationalisierung eines Confounders nur einen begrenzten zurückliegenden Zeitraum umfasst (z. B. Anzahl Schübe im vorangegangenen Jahr anstelle von Anzahl Schübe in den vorangegangenen 3 Jahren). Ein weiteres Beispiel stellt die Operationalisierung von Komorbiditäten dar. Im vorliegenden Beispiel der RRMS wurden eine ganze Reihe von Komorbiditäten als potenzielle Confounder identifiziert. Werden diese übergeordnet als 1 potenzieller Confounder „Komorbiditäten“ zusammengefasst, geht dies mit dem größten Informationsverlust einher. Eine Alternative dazu wäre es, die wichtigsten Komorbiditäten herauszugreifen, oder aber jede einzelne Komorbidität separat als Confounder zu betrachten. Je nach zur Verfügung stehender Patientenzahl ist abzuwägen, ob der durch die Zusammenfassung in Kauf genommene Informationsverlust vertretbar ist.

Nicht messbare bzw. nicht operationalisierbare potenzielle Confounder können zwar nicht in die Auswertung einbezogen werden, sollten jedoch benannt und bei der Ergebnisinterpretation berücksichtigt werden.

Es gibt potenzielle Confounder, wie z. B. den Patientenwunsch oder die Arztentscheidung, die aufgrund ihrer Komplexität nicht adäquat erhoben bzw. nicht operationalisiert werden können. Diese potenziellen Confounder sollten dennoch klar benannt werden und es sollte begründet werden, warum eine Operationalisierung nicht möglich ist. Das gilt ebenfalls für potenzielle Confounder, die auf komplexen Untersuchungsmethoden basieren, welche nicht regelhaft im Versorgungsalltag durchgeführt werden, und diese dadurch nicht erhoben werden (z. B. regelmäßige Untersuchung der Cerebrospinalflüssigkeit außerhalb der MS-Diagnose).

Diese potenziellen Confounder können zwar nicht in die Auswertung einbezogen werden, dies muss jedoch bei der Ergebnisinterpretation berücksichtigt werden [10].

5.3.3 Schritt 3: Relevanzbewertung

Die Relevanzbewertung der potenziellen Confounder durch klinische Expertinnen und Experten sollte als 3. Schritt erfolgen. Eine Bewertung mittels Punktevergabe wird empfohlen.

Grundsätzlich gibt es verschiedene Herangehensweisen die Relevanzbewertung der potenziellen Confounder durchzuführen, wie z. B. ein Ranking der 5 wichtigsten Confounder oder eine Bewertung mittels Punktevergabe durch die klinischen Expertinnen und Experten, wie stark jede einzelne identifizierte Confounder-Variable die Therapieentscheidung beeinflusst (z. B. auf einer 5-Punkte-Skala von 1 [keine Berücksichtigung] bis 5 [starke Berücksichtigung]). Der Vorteil der Punktevergabe ist, dass jeder einzelne potenzielle Confounder einen Score erhält und der Informationsgehalt im Vergleich zur Benennung der 5 wichtigsten Confounder deutlich höher ist. Da diese Art der Bewertung aufwendiger ist, wird empfohlen, bereits vorher die identifizierten Confounder so weit möglich zusammenzuführen, um den Aufwand zu reduzieren.

Die vergebenen Scores sollten nicht zum Ausschluss potenzieller Confounder genutzt werden. Alle identifizierten potenziellen Confounder sollten erhoben werden. Das Ergebnis des Bewertungsschritts kann aber in der Datenanalyse für eine inhaltlich gestützte Variablenselektion verwendet werden in solchen Fällen, in denen das statistische Modell nicht konvergiert. Basierend auf dem Score kann dann entschieden werden, welcher Confounder nicht in das Modell einbezogen wird. Der potenzielle Confounder mit dem niedrigsten Score wird beispielsweise als erstes nicht mehr im Modell berücksichtigt.

Die in diesem Schritt eingebundene Gruppe klinischer Expertinnen und Experten sollten im Idealfall nicht ausschließlich aus den in Schritt 1 zur Identifikation involvierten klinischen Expertinnen und Experten bestehen. Auch hier ist es erforderlich, die klinischen Expertinnen und Experten namentlich zu benennen und potenzielle Interessenkonflikte offenzulegen.

6 Diskussion

In nicht randomisierten Vergleichen von Interventionen ist für eine adäquate Confounderkontrolle eine systematische Identifikation der potenziellen Confounder Voraussetzung. Das Ziel der Confounderkontrolle ist es trotz der fehlenden Randomisierung eine möglichst niedrig verzerrte Schätzung des interessierenden kausalen Effekts zu erhalten. Solche nicht randomisierten Studiendesigns spielen u. a. vor dem Hintergrund der anwendungs-
begleitenden Datenerhebungen gemäß § 35a Abs. 3b SGB V, die gemäß gesetzlicher Vorgaben ohne Randomisierung durchzuführen sind, in der Nutzenbewertung eine immer größere Rolle. Aber auch in Situationen, in denen für die zu bewertende Intervention lediglich 1-armige Studien vorliegen und ein Vergleich einzelner Arme verschiedener Studien oder ein Vergleich mit einer externen Kontrolle angestrebt wird, spielt eine adäquate Identifikation von und Adjustierung für potenzielle Confounder eine zentrale Rolle.

Trotz dieser Zahl von nicht randomisierten vergleichenden Studien stellt die Publikation von Pufulete 2022 [11] bisher den einzigen publizierten Vorschlag einer systematischen Confounderidentifikation dar. Auch über eine fokussierte Informationsbeschaffung wurden keine weiteren Publikationen zu systematischen Ansätzen identifiziert.

Pufulete 2022 schlagen für die systematische Confounderidentifikation 3 methodische Bausteine vor: eine systematische Literaturrecherche, Interviews mit Klinikerinnen und Klinikern, und eine Umfrage unter Klinikerinnen und Klinikern. Die vorliegenden Untersuchungen in der Indikation der RRMS haben die Relevanz der 3 Bausteine bestätigt. So wären durch einen rein literaturgestützten Ansatz nicht alle wichtigen bekannten Confounder (wie bspw. patientenzentrierte Faktoren) identifiziert worden. Die teils mangelhafte Berichterstattung von potenziellen Confoundern, welche identifiziert und im statistischen Modell berücksichtigt werden, stellt ein bestehendes Problem dar [34]. Es ist davon auszugehen, dass die Interviews mit klinischen Expertinnen und Experten eine noch größere Bedeutung haben, wenn in der vorhandenen Literatur nur eine geringe Anzahl an Studien vorliegt, die für die Identifikation von möglichen Confoundern herangezogen werden kann. Dies ist erwartbar für eine Reihe von Indikationen der anwendungs-
begleitenden Datenerhebung der Fall, die insbesondere in seltenen Erkrankungen durchgeführt wird.

Der Ansatz von Pufulete 2022, der neben der systematischen Literaturrecherche auch die Einbindung von klinischen Expertinnen und Experten vorsieht, wird von den Autoren selbst als sehr ressourcenintensiv bewertet. Dies konnte von den externen Sachverständigen hinsichtlich der Confounderidentifikation in der Indikation RRMS bestätigt werden. Für die Identifikation der 160 potenziellen Confounder-Variablen waren 62 Personentage notwendig. Nichtsdestotrotz ist das Verfahren grundsätzlich machbar. Der größte Zeitbedarf, aber auch die höchste Zahl an identifizierten Confoundern war der systematischen Literaturrecherche zuzuschreiben. Inwieweit der hohe Zeitaufwand auf andere Indikationen übertragbar ist, ist

unklar. Sowohl bei Pufulete 2022 als auch in der vorliegenden Untersuchung waren Indikationen mit sehr umfangreicher Studienlage Gegenstand der Untersuchungen.

Unter den 160 von den externen Sachverständigen identifizierten Variablen zeigten sich eine Reihe von inhaltlich ähnlichen oder sich überlappenden Konzepten. In der vorliegenden Untersuchung konnten die 160 identifizierten Variablen auf 28 potenzielle Confounder zusammengeführt werden. Die weitere Zusammenführung der identifizierten potenziellen Confounder-Variablen stellt somit einen zentralen Schritt im Sinne der Machbarkeit dar, nicht nur für die Datenerhebung sondern auch für die statistischen Analysen.

Auch wenn die Zusammenführung erfolgreich war, ist in der Regel dennoch nicht davon auszugehen, dass in einer bestehenden Datenquelle (z. B. einem Register) bereits eine vollumfängliche Erhebung aller potenziellen Confounder erfolgt. Insbesondere bei einem Vergleich einzelner Arme verschiedener Studien oder mit einer externen Kontrolle wird auf bereits erhobene retrospektive Daten zurückgegriffen. In diesen Datenquellen liegen häufig keine Informationen zu den identifizierten potenziellen Confoundern vor, bzw. wenn Informationen vorliegen, sind diese häufig unvollständig [22,35-37]. Eine retrospektive Datenerhebung kommt somit häufig nicht infrage. Dies entspricht den bisherigen Erfahrungen aus der anwendungsbegleitenden Datenerhebung, in der in bisher allen Fällen eine prospektive Datenerhebung notwendig ist, um die notwendige Erhebung der Confounder sicherzustellen.

Das mit dem Bericht der externen Sachverständigen und den eigenen Untersuchungen im vorliegenden Arbeitspapier identifizierte Confounder-Set soll zur Nachbildung einer konkreten RCT genutzt werden. Eine solche Nachbildung einer RCT kann in der Folge z. B. dazu dienen, Übertragbarkeitsfragestellungen (z. B., ob die in einer RCT beobachteten Effekte bei jüngeren Erwachsenen [18 bis 55 Jahre] auf ältere Erwachsene [> 55 Jahre] übertragbar sind) zu beantworten. War z. B. die Nachbildung der Ergebnisse einer RCT bei jüngeren Erwachsenen mittels Registerdaten erfolgreich, können im Anschluss diese Registerdaten zur Untersuchung der Fragestellung bei Älteren herangezogen werden. Ist eine solche Fragestellung Anlass der Confounderidentifikation, sollte dies bereits bei der Definition der Einschlusskriterien der Literaturrecherche berücksichtigt werden. So sollten die Einschlusskriterien beide Populationen abbilden (im genannten Beispiel Erwachsene und Kinder und Jugendliche). Sollten einzelne der identifizierten potenziellen Confounder nur für eine der beiden Fragestellungen Relevanz haben, können diese für Teilfragestellungen begründet ausgeschlossen bzw. nur für eine der Fragestellungen aufgenommen werden.

Ob eine bestimmte Variable einen Confounder darstellt, ist eine Frage der Kausalität und lässt sich in den zu untersuchenden Daten nicht statistisch testen [38]. In bestimmten Studiensituationen können protokollbedingt auch rein prognostische Faktoren zu Confoundern werden [39]. Ebenso liegen meist nur unzureichend Informationen vor, um

zuverlässig die kausalen Zusammenhänge zwischen den verschiedenen Variablen zu spezifizieren, die die Therapieentscheidung und das Behandlungsergebnis bestimmen [38]. Ähnlich beschreiben Pufulete 2022, dass es keine Sicherheit dafür gibt, dass die von ihnen ausgewählten Variablen ausreichen, um eine Verzerrung durch Confounder zu kontrollieren. Entsprechend unterliegen auch die im vorliegenden Bericht getroffenen Entscheidungen insbesondere in Bezug auf die vorgenommene Zusammenführung oder die vorgeschlagenen Operationalisierungen lediglich Annahmen, die nicht überprüfbar sind. Ob es sich bei den identifizierten potenziellen Confoundern ggf. um Mediatoren oder Collider handelt, wurde nicht überprüft. Der Ausschluss solcher Variablen sollte literaturgestützt begründet werden und die wechselseitigen Beziehungen der potenziellen Confounder mithilfe eines gerichteten azyklischen Graphen (DAG) [1,2] visualisiert werden.

Selbst bei sorgfältiger systematischer Identifikation von Confoundern, deren Erhebung in hoher Qualität und deren vollständiger Berücksichtigung in der Analyse, bleibt aufgrund von nicht bekannten Confoundern bzw. weiterer nicht berücksichtigter oder nicht messbarer Confounder die Gefahr von residuellem / unmeasured Confounding. Diese Unsicherheiten müssen bei der Ergebnisinterpretation berücksichtigt werden. Eine Möglichkeit stellt die Anwendung einer verschobenen Nullhypothese dar, wie sie für die anwendungsbegleitende Datenerhebung vorgeschlagen wurde [10]. Ein (positiver oder negativer) Effekt wird erst dann abgeleitet, wenn das Konfidenzintervall für den beobachteten Effekt oberhalb einer zu definierenden Schwelle liegt.

Zusammenfassend wird das im vorliegenden Arbeitspapier vorgeschlagene Vorgehen zur Identifikation als durchführbar angesehen. Die vorliegenden Untersuchungen unterstreichen jedoch auch, wie aufwendig die Confounderidentifikation für nicht randomisierte vergleichende Studien ist. In einem nächsten Schritt soll deshalb geprüft werden, welche weiteren möglichen Optionen zur Aufwandsreduktion bestehen. Dies betrifft zum einen Möglichkeiten den Aufwand für die Confounderidentifikation zu reduzieren (z. B. Reduktion des Recherche-aufwands). Zum anderen betrifft dies Möglichkeiten, die Anzahl der potenziellen Confounder auf Basis ihrer inhaltlichen Bedeutung zu verringern (z. B. anhand ihres möglichen Einflusses auf die Effektschätzungen) und damit den Aufwand für die Erhebung zu reduzieren. Die möglichen Optionen zur Aufwandsreduktion müssen dem damit einhergehenden Informationsverlust sowie dem potenziellen Verlust der Ergebnissicherheit gegenübergestellt werden.

Die einzige Möglichkeit, die erhöhten Anforderungen an den Umfang der Datenerhebung durch die erforderliche Identifikation und Erhebung potenzieller Confounder zu reduzieren und die bestehende Restunsicherheit bezüglich z. B. nicht messbarer Confounder und deren fehlende Berücksichtigung in der Auswertung zu vermeiden, ist die Randomisierung. Die Durchführung einer randomisierten kontrollierten Studie erspart zum einen die

Notwendigkeit die potenziellen Confounder zu identifizieren und zu erheben, zum anderen wird die erforderliche Anzahl der Patientinnen und Patienten deutlich reduziert und letztlich die Machbarkeit erhöht. Gleichzeitig ergibt sich eine deutlich höherer Ergebnissicherheit. RCTs sind auch bei seltenen Erkrankungen grundsätzlich durchführbar [40]. Die häufige Kritik, dass RCTs keine gute externe Validität aufweisen, ließe sich durch breite Ein- und Ausschlusskriterien im Sinne einer pragmatischen RCT adressieren [41-43]. Innerhalb der RCTs stellen registerbasierte RCTs eine inhaltlich sinnvolle Möglichkeit dar, weil mit der Verwendung einer bestehenden Datenstruktur ein weiterer Effizienzgewinn einhergeht [44].

7 Fazit

Die externen Sachverständigen haben unter Anwendung des von Pufulete 2022 vorgeschlagenen Vorgehens insgesamt 160 potenzielle Confounder-Variablen identifiziert. Diese wurden in weiterführenden eigenen Untersuchungen zu 28 potenziellen Confoundern zusammengeführt. Damit konnte demonstriert werden, dass die Confounderidentifikation gemäß Pufulete 2022 grundsätzlich machbar ist, z. B. im Rahmen einer anwendungs-
begleitenden Datenerhebung.

Aus den in den Untersuchungen gewonnenen Erfahrungen wurden die folgenden Handlungsempfehlungen für eine systematische Confounderidentifikation in 3 Schritten abgeleitet:

1) Confounderidentifikation

Potenzielle Confounder-Variablen sollen aus einer systematischen Literaturrecherche und aus Interviews mit klinischen Expertinnen und Experten identifiziert werden.

2) Confounder-Zusammenführung

Confounder-Variablen, welche ein inhaltlich gleiches Konstrukt messen oder sich inhaltlich überlappen, sollen zusammengeführt werden. Für jeden dieser potenziellen Confounder ist eine eindeutige Operationalisierung festzulegen.

3) Confounder-Bewertung

Jeder potenzielle Confounder sollte über eine Umfrage mit klinischen Expertinnen und Experten in seiner Bedeutung bewertet werden. Das Ergebnis kann für eine inhaltlich gestützte Variablenselektion verwendet werden.

Insbesondere der 2. Schritt stellt dabei eine Weiterentwicklung der von Pufulete 2022 vorgeschlagenen Methodik dar, mit dem Ziel, den Aufwand ohne relevanten Informationsverlust zu reduzieren.

Das Verfahren der systematischen Confounderidentifikation ist trotz der beschriebenen Möglichkeiten zur Aufwandseinsparung ressourcenintensiv. In einem nächsten Schritt soll geprüft werden, welche weiteren möglichen Optionen zur Aufwandsreduktion bestehen, sowohl für die Confounderidentifikation als auch die zugehörige Erhebung der potenziellen Confounder. Die möglichen Optionen zur Aufwandsreduktion müssen dem damit einhergehenden Informationsverlust sowie dem potenziellen Verlust der Ergebnissicherheit gegenübergestellt werden.

Um in einer nicht randomisierten vergleichenden Studie die Verzerrung durch Confounding möglichst zu minimieren, ist eine systematische Confounderidentifikation erforderlich. Dennoch verbleibt in jedem Fall eine Unsicherheit, ob wirklich alle relevanten Confounder

identifiziert wurden und dementsprechend in der Analyse berücksichtigt werden können. Nicht bekannte Confounder ebenso wie nicht messbare Confounder bleiben ein ungelöstes Problem, sodass in nicht randomisierten vergleichenden Studien nur bei hinreichend großen Effekten bspw. ein Zusatznutzen einer Therapie ableitbar ist. Eine sinnvolle Alternative ist einzig die Durchführung einer randomisierten vergleichenden Studie, wie z. B. einer registerbasierten RCT, für die keine Confounderidentifikation und -kontrolle erforderlich ist.

8 Literatur

1. Feeney T, Pires Hartwig F, Davies NM. How to use directed acyclic graphs: guide for clinical researchers. *BMJ* 2025; 388: e078226. <https://doi.org/10.1136/bmj-2023-078226>.
2. Williams TC, Bach CC, Matthiesen NB et al. Directed acyclic graphs: a tool for causal studies in paediatrics. *Pediatr Res* 2018; 84(4): 487-493. <https://doi.org/10.1038/s41390-018-0071-3>.
3. Catalogue of bias collaboration, Lee H, Aronson JK et al. Catalogue Of Bias; Collider bias [online]. 2019 [Zugriff: 25.02.2025]. URL: <https://catalogofbias.org/biases/collider-bias/>.
4. Sendor R, Sturmer T. Core concepts in pharmacoepidemiology: Confounding by indication and the role of active comparators. *Pharmacoepidemiol Drug Saf* 2022; 31(3): 261-269. <https://doi.org/10.1002/pds.5407>.
5. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol* 2016; 183(8): 758-764. <https://doi.org/10.1093/aje/kwv254>.
6. Desai RJ, Wang SV, Sreedhara SK et al. Process guide for inferential studies using healthcare data from routine clinical practice to evaluate causal effects of drugs (PRINCIPLED): considerations from the FDA Sentinel Innovation Center. *BMJ* 2024; 384: e076460. <https://doi.org/10.1136/bmj-2023-076460>.
7. Hernan MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *JAMA* 2022; 328(24): 2446-2447. <https://doi.org/10.1001/jama.2022.21383>.
8. Wang SV, Schneeweiss S, RCT-DUPLICATE Initiative. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. *JAMA* 2023; 329(16): 1376-1385. <https://doi.org/10.1001/jama.2023.4221>.
9. Mathes T. Anforderung an die Daten für die Target-Trial-Emulation: Eine Diskussion unter Betrachtung von Patientenregistern. *GMS Med Inform Biom Epidemiol* 2024; 20: Doc03. <https://doi.org/10.3205/mibe000259>.
10. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach § 35a SGB V; Rapid Report [online]. 2020 [Zugriff: 02.06.2020]. URL: https://www.iqwig.de/download/A19-43_Versorgungsnaher-Daten-zum-Zwecke-der-Nutzenbewertung_Rapid-Report_V1-1.pdf.
11. Pufulete M, Mahadevan K, Johnson TW et al. Confounders and co-interventions identified in non-randomized studies of interventions. *J Clin Epidemiol* 2022; 148: 115-123. <https://doi.org/10.1016/j.jclinepi.2022.03.018>.

12. Reeves BC, Deeks JJ, Higgins JPT et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.5; Chapter 24: Including non-randomized studies on intervention effects [last updated October 2019] [online]. 2024 [Zugriff: 12.03.2025]. URL: <https://training.cochrane.org/handbook/current/chapter-24>.
13. Sterne JAC, Hernán MA, McAleenan A et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.5; Chapter 25: Assessing risk of bias in a non-randomized study [last updated October 2019] [online]. 2024 [Zugriff: 13.03.2025]. URL: <https://training.cochrane.org/handbook/current/chapter-25>.
14. Vanier A, Fernandez J, Kelley S et al. Rapid access to innovative medicinal products while ensuring relevant health technology assessment. Position of the French National Authority for Health. *BMJ Evid Based Med* 2024; 29(1): 1-5. <https://doi.org/10.1136/bmjebm-2022-112091>.
15. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden; Version 7.0 [online]. 2023 [Zugriff: 02.09.2024]. URL: https://www.iqwig.de/methoden/allgemeine-methoden_version-7-0.pdf.
16. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Exagamglogen Autotemcel (Sichelzellkrankheit); Bewertung gemäß § 35a SGB V; AbD-Konzept [online]. 2023 [Zugriff: 23.01.2025]. URL: https://www.iqwig.de/download/a23-49_exagamglogen-autotemcel_abd-konzept_v1-0.pdf.
17. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Etranacogen Dezaparovec (Hämophilie B) – Bewertung gemäß § 35a SGB V; AbD-Konzept [online]. 2023 [Zugriff: 11.07.2023]. URL: https://www.iqwig.de/download/a22-83_etranacogen-dezaparovec-haemophilie-b_abd-konzept_v1-1.pdf.
18. CSL Behring. Study Protocol; Routine Practice Data Collection and Evaluation of etranacogene dezaparovec (Hemgenix®) and prophylactic factor IX (FIX) replacement in severe and moderately severe haemophilia B without a history of FIX inhibitors: a prospective, non-interventional study mandated by G-BA [online]. 2024 [Zugriff: 24.03.2025]. URL: https://www.g-ba.de/downloads/40-268-10667/2024-07-18_AM-RL-XII_Etranacogen-Dezaparovec_2022-AbD-005_Studienunterlagen.pdf.
19. Novartis Gene Therapies. Routine data collection and evaluations of onasemnogene abeparovec in Germany, Study Protocol [online]. 2024 [Zugriff: 24.03.2025]. URL: https://www.g-ba.de/downloads/40-268-10961/2024-06-06_AM-RL-XII_Onasemnogen-Abeparovec_2020-AbD-001_Ueberpruefung-SP-SAP_Studienunterlagen.pdf.

20. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen.
Anwendungsbegleitende Datenerhebung zu Onasemnogen-Abeparvovec: Prüfung des Studienprotokolls und des statistischen Analyseplans; Addendum zum Auftrag A20-61 [online]. 2021 [Zugriff: 17.03.2025]. URL: https://www.iqwig.de/download/a21-107_anwendungsbegleitende-datenerhebung-onasemnogen-abeparvovec-pruefung-protokoll-und-sap_addendum-zum-auftrag-a20-61_v1-0.pdf.
21. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen.
Anwendungsbegleitende Datenerhebung zu Brexucabtagen Autoleucel: Prüfung des Studienprotokolls (Version 1.0) und des statistischen Analyseplans (Version 1.0); 2. Addendum zum Projekt A21-130 [online]. 2023 [Zugriff: 17.03.2025]. URL: https://www.iqwig.de/download/a22-136_brexucabtagen-autoleucel_addendum-zum-projekt-a21-130_v1-0.pdf.
22. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen.
Anwendungsbegleitende Datenerhebung zu Etranacogen Dezaparvovec: Prüfung des Studienprotokolls (Version 1.0) und des statistischen Analyseplans (Version 1.0); 2. Addendum zum Projekt A22-83 [online]. 2023 [Zugriff: 17.03.2025]. URL: <https://doi.org/10.60584/A23-99>.
23. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen.
Anwendungsbegleitende Datenerhebung zu Risdiplam: Prüfung des Studienprotokolls (Version 3.0) und des statistischen Analyseplans (Version 3.0); 3. Addendum zum Projekt A21-131 [online]. 2024 [Zugriff: 17.03.2025]. URL: <https://doi.org/10.60584/A24-57>.
24. Fox RJ, Miller DH, Phillips JT et al. Placebo-controlled phase 3 study of oral BG-12 or glatiramer in multiple sclerosis. *N Engl J Med* 2012; 367(12): 1087-1097. <https://doi.org/10.1056/NEJMoa1206328>.
25. Schneeweiss S. Automated data-adaptive analytics for electronic healthcare data to study causal treatment effects. *Clin Epidemiol* 2018; 10: 771-788. <https://doi.org/10.2147/CLEP.S166545>.
26. Zeng J, Gensheimer MF, Rubin DL et al. Uncovering interpretable potential confounders in electronic medical records. *Nat Commun* 2022; 13(1): 1014. <https://doi.org/10.1038/s41467-022-28546-8>.
27. Harris J, Pouwels KB, Johnson T et al. Bleeding risk in patients prescribed dual antiplatelet therapy and triple therapy after coronary interventions: the ADAPTT retrospective population-based cohort studies. *Health Technol Assess* 2023; 27(8): 1-257. <https://doi.org/10.3310/MNJY9014>.
28. VanderWeele TJ. Principles of confounder selection. *Eur J Epidemiol* 2019; 34(3): 211-219. <https://doi.org/10.1007/s10654-019-00494-6>.

29. Hirt J, Hemkens LG. Systematische Confounderidentifikation in der Indikation schubförmig remittierende multiple Sklerose (RRMS); Supplement (Anhang F) zur externen Expertise zum IQWiG-Arbeitspapier GA23-02. [Demnächst verfügbar unter: https://www.iqwig.de/download/ga23-02_supplement-zur-externen-expertise_anhang-f.xlsx].
30. Viswanathan M, Berkman ND, Dryden DM et al. Assessing Risk of Bias and Confounding in Observational Studies of Interventions or Exposures: Further Development of the RTI Item Bank [online]. 2013 [Zugriff: 13.03.2025]. URL: <https://www.ncbi.nlm.nih.gov/books/n/methresbias/pdf/>.
31. Harrell FE Jr. Regression Modeling Strategies; with Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. New York: Springer; 2001.
32. Peduzzi P, Concato J, Kemper E et al. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996; 49(12): 1373-1379. [https://doi.org/10.1016/s0895-4356\(96\)00236-3](https://doi.org/10.1016/s0895-4356(96)00236-3).
33. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: A systematic review of empirical tests. Soc Sci Med 2022; 292: 114523. <https://doi.org/10.1016/j.socscimed.2021.114523>.
34. Petersen JM, Barrett M, Ahrens KA et al. The confounder matrix: A tool to assess confounding bias in systematic reviews of observational studies of etiology. Res Synth Methods 2022; 13(2): 242-254. <https://doi.org/10.1002/jrsm.1544>.
35. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Amivantamab (NSCLC) – Nutzenbewertung gemäß § 35a SGB V; Dossierbewertung [online]. 2022 [Zugriff: 17.03.2025]. URL: https://www.iqwig.de/download/a22-05_amivantamab_nutzenbewertung-35a-sgb-v_v1-0.pdf.
36. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Ravulizumab (Neuromyelitis-optica-Spektrum-Erkrankungen); Nutzenbewertung gemäß § 35a SGB V; Dossierbewertung [online]. 2023 [Zugriff: 17.03.2025]. URL: https://www.iqwig.de/download/a23-50_ravulizumab_nutzenbewertung-35a-sgb-v_v1-0.pdf.
37. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Lanadelumab (hereditäres Angioödem) – Nutzenbewertung gemäß § 35a SGB V; Dossierbewertung [online]. 2021 [Zugriff: 17.03.2025]. URL: https://www.iqwig.de/download/a21-63_lanadelumab_nutzenbewertung-35a-sgb-v_v1-0.pdf.
38. Brookhart MA, Sturmer T, Glynn RJ et al. Confounding control in healthcare database research: challenges and potential approaches. Med Care 2010; 48(6 Suppl): S114-120. <https://doi.org/10.1097/MLR.0b013e3181d8be3>.

39. Brookhart MA, Schneeweiss S, Rothman KJ et al. Variable selection for propensity score models. *Am J Epidemiol* 2006; 163(12): 1149-1156. <https://doi.org/10.1093/aje/kwj149>.
40. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Evidenz zu Orphan Drugs; Arbeitspapier [online]. 2021 [Zugriff: 11.07.2023]. URL: https://www.iqwig.de/download/ga21-01_evidenz-zu-orphan-drugs_arbeitspapier_v1-0.pdf.
41. Mathes T, Buehn S, Prengel P et al. Registry-based randomized controlled trials merged the strength of randomized controlled trails and observational studies and give rise to more pragmatic trials. *J Clin Epidemiol* 2018; 93: 120-127. <https://doi.org/10.1016/j.jclinepi.2017.09.017>.
42. Lange S, Lauterberg J. Pragmatischere randomisierte Studien mit Fokus auf Registerbasierung. *Präv Gesundheitsf* 2022. <https://doi.org/10.1007/s11553-022-00974-w>.
43. Ford I, Norrie J. Pragmatic Trials. *N Engl J Med* 2016; 375(5): 454-463. <https://doi.org/10.1056/NEJMra1510059>.
44. Lauer MS, D'Agostino RB Sr. The randomized registry trial -the next disruptive technology in clinical research? *N Engl J Med* 2013; 369(17): 1579-1581. <https://doi.org/10.1056/NEJMp1310102>.

Anhang A Informationsbeschaffung

A.1 Liste der ausgeschlossenen Publikationen aus der fokussierten Informationsbeschaffung zur systematischen Vorgehensweise der Confounderidentifikation

1. Cobzaru R, Jiang S, Ng K et al. State of the Art Causal Inference in the Presence of Extraneous Covariates: A Simulation Study. AMIA Annual Symposium Proceedings/AMIA Symposium 2021; 2021: 334-342.
2. Huang B, Szczesniak R. Can't see the wood for the trees: confounders, colliders and causal inference - a statistician's approach. Thorax 2019; 74(4): 323-325.
<https://doi.org/10.1136/thoraxjnl-2018-212780>.
3. Inoue K, Sakamaki K, Komukai S et al. Confounder Selection and Sensitivity Analyses to Unmeasured Confounding from Epidemiological and Statistical Perspectives. J Epidemiol 2024. <https://doi.org/10.2188/jea.JE20240082>.
4. Keen R, Tiemeier H. Covariate Selection from Data Collection Onwards: A Methodology for Neurosurgeons. World Neurosurg 2022; 161: 245-250.
<https://doi.org/10.1016/j.wneu.2021.11.057>.
5. Luijken K, Groenwold RHH, van Smeden M et al. A comparison of full model specification and backward elimination of potential confounders when estimating marginal and conditional causal effects on binary outcomes from observational data. Biom J 2024; 66(1): e2100237. <https://doi.org/10.1002/bimj.202100237>.
6. Malec SA, Wei P, Bernstam EV et al. Using computable knowledge mined from the literature to elucidate confounders for EHR-based pharmacovigilance. J Biomed Inform 2021; 117: 103719. <https://doi.org/10.1016/j.jbi.2021.103719>.
7. Petersen JM, Barrett M, Ahrens KA et al. The confounder matrix: A tool to assess confounding bias in systematic reviews of observational studies of etiology. Res Syn Meth 2022; 13(2): 242-254. <https://doi.org/10.1002/jrsm.1544>.
8. Quint JK, Minelli C. Can't see the wood for the trees: confounders, colliders and causal inference - a clinician's approach. Thorax 2019; 74(4): 321-322.
<https://doi.org/10.1136/thoraxjnl-2018-212488>.
9. Steiner PM, Cook TD, Shadish WR et al. The importance of covariate selection in controlling for selection bias in observational studies. Psychol Methods 2010; 15(3): 250-267.
<https://doi.org/10.1037/a0018719>.
10. Tang W, Spiegelman D, Liao X et al. Causal Selection of Covariates in Regression Calibration for Mismeasured Continuous Exposure. Epidemiology 2024; 35(3): 320-328.
<https://doi.org/10.1097/ede.0000000000001706>.

11. Tanner-Smith EE, Lipsey MW. Identifying Baseline Covariates for Use in Propensity Scores: A Novel Approach Illustrated for a Non-randomized Study of Recovery High Schools. *Peabody Journal of Education* 2014; 89(2): 183-196.

<https://doi.org/10.1080/0161956x.2014.895647>.

12. van Zwieten A, Tennant PWG, Kelly-Irving M et al. Avoiding overadjustment bias in social epidemiology through appropriate covariate selection: a primer. *J Clin Epidemiol* 2022; 149: 127-136. <https://doi.org/10.1016/j.jclinepi.2022.05.021>.

13. Zawadzki RS, Grill JD, Gillen DL. Frameworks for estimating causal effects in observational settings: comparing confounder adjustment and instrumental variables. *BMC Med Res Methodol* 2023; 23(1): 122. <https://doi.org/10.1186/s12874-023-01936-2>.

14. Zeng J, Gensheimer MF, Rubin DL et al. Uncovering interpretable potential confounders in electronic medical records. *Nature communications* 2022; 13(1): 1014.

<https://doi.org/10.1038/s41467-022-28546-8>.

A.2 Suchstrategien

A.2.1 Bibliografische Datenbanken

1. MEDLINE

Suchoberfläche: Ovid

Ovid MEDLINE(R) ALL 1946 to December 12, 2024

#	Searches
1	(confounder* or covariate*).ti.
2	((covariate\$1 or confounder\$1) adj3 (identif* or select* or bias or control)).ti,ab.
3	(causal adj (inference* or effect*)).ti,ab.
4	(target* adj3 trial* adj3 emulat*).ti,ab.
5	3 or 4
6	((observational* or cohort or non-randomized) adj3 (stud* or data)).ti,ab.
7	and/1-2,6
8	and/1,5
9	or/7-8

Anhang B Zusammenführung der identifizierten potenziellen Confounder

Tabelle 2: Zuordnung der identifizierten potenziellen Confounder-Variablen zu den jeweiligen potenziellen Confoundern (mehreseitige Tabelle)

Nr. ^a	potenzieller Confounder ^b	Nr. ^c	identifizierte potenzielle Confounder-Variablen ^b	n (Lit ^d)	Int ^e	Score ^f
1	gender/sex	1	gender/sex	61	ja	2,43
2	age	2	age	70	ja	3,86
3	EDSS score	3	EDSS score (mean or median)	61	nein	2,88
		4	Disability status	1	ja	3,25
		5	Neurological symptoms	0	ja	3,5
		6	Cognition (self-rated)	1	nein	3,43
		7	Functional System (FS) score	1	nein	2,29
		8	Numbness	1	nein	2,57
		9	Patient Determined Disease Steps (PDDS) score	2	nein	2,29
4	EDSS change preceding year	10	EDSS change preceding year	1	nein	3,88
		11	Deterioration index (defined as the EDSS score divided by the disease duration)	2	nein	4
		12	EDSS in the past 3 months	1	nein	3
		13	MS Severity Score (MSSS)	1	nein	1,5
		14	Disease activity	2	nein	4,5
		15	Disease course	1	nein	na
		16	Remit of relapse	0	ja	4,43
17	Time since last relapse	5	nein	3,71		
5	MS Functional Composite (MSFC) score	18	MS Functional Composite (MSFC) score	3	nein	1,75
		19	Ambulation index	3	nein	2,88
		20	Clinical condition	0	ja	4,25
6	Fatigue	21	Fatigue	1	nein	2,86
		22	Modified Fatigue Impact Scale (MFIS) score	1	nein	2,14

Tabelle 2: Zuordnung der identifizierten potenziellen Confounder-Variablen zu den jeweiligen potenziellen Confoundern (mehreseitige Tabelle)

Nr. ^a	potenzieller Confounder ^b	Nr. ^c	identifizierte potenzielle Confounder-Variablen ^b	n (Lit ^d)	Int ^e	Score ^f
7	Comorbidities	23	Comorbidities (general)	2	ja	4
		24	Depression diagnosis	2	nein	2
		25	Hamilton Anxiety Rating Scale (HAM-A)	1	nein	na
		26	Hamilton Depression Rating Scale (HAM-D)	1	nein	na
		27	Allergy	1	nein	na
		28	Cardiovascular comorbidities	1	ja	2,5
		29	Charlson Comorbidity Index (CCI) score	3	nein	1,86
		30	Chronic infections	0	ja	4,13
		31	Dermatological diseases	0	ja	2,88
		32	Diabetes	1	nein	1,75
		33	Headache	1	nein	na
		34	Skin problems	0	ja	2,25
		35	Thyroid dysfunction	1	nein	na
		36	Tuberculosis	0	ja	2,38
		37	Bowel symptoms	1	nein	3,86
		38	Depression (self-rated)	1	nein	3
		39	Gastrointestinal disorders	0	ja	4
		40	Hospitalisations in previous year	1	nein	2
8	Absolute lymphocyte count; Lymphopenia	41	Absolute lymphocyte count; Lymphopenia	1	ja	4,57
9	Study center; MS center	42	Study center; MS center	3	nein	na
		43	Country	6	nein	1,14
		44	Region	5	nein	1,14
		45	Within-country region	2	nein	1,14
		46	Type and speciality of prescriber	2	nein	2,71
10	BMI	47	BMI; or height and weight	6	nein	2,29
		48	Overweight or obesity	1	nein	2,29
		49	Weight	2	nein	2
11	Tobacco use	50	Tobacco use	1	nein	2,14
12	infratentorial lesion	51	Localisation of lesions	0	ja	3,57
		52	Pyramidal impairment	1	ja	3,38
		53	Spinal lesions	0	ja	4,14

Tabelle 2: Zuordnung der identifizierten potenziellen Confounder-Variablen zu den jeweiligen potenziellen Confoundern (mehreseitige Tabelle)

Nr. ^a	potenzieller Confounder ^b	Nr. ^c	identifizierte potenzielle Confounder-Variablen ^b	n (Lit ^d)	Int ^e	Score ^f
13	Number of contrast-enhancing lesions	54	Number of gadolinium enhancing lesions; Number of contrast-enhancing lesions; Number of patients with gadolinium enhancing lesions; Patients with enhancement	21	ja	4,57
		55	Status of gadolinium enhancing T1 activity	1	nein	4,29
14	T1 hypointense lesion volume	56	T1 hypointense lesion volume	7	nein	3
15	Number of T2 lesions	57	Number of T2 lesions	12	nein	4,14
		58	MRI status	1	ja	3
		59	MRI brain lesions; number of lesions	2	ja	3,86
		60	New lesions	0	ja	4,29
		61	Presence of at least MRI T2 lesions and at least one contrast-enhancing MRI lesion	2	nein	3,43
16	Volume of T2 lesions	62	Volume of T2 lesions	7	nein	3,14
		63	Enhancing lesion volume	3	nein	3
		64	Lesion load	0	ja	3,71
		65	Size of lesions	0	ja	3,14
		66	Brain volume; normalized brain volume	7	ja	2,17
		67	Normalized grey matter volume	1	nein	2,29
		68	Normalized white matter volume	1	nein	2,43
17	Tolerance of previous therapy and expectations of the treatment	69	Reason for discontinuing prior therapy	2	nein	4,67
		70	Drug efficacy (known or suspected)	0	ja	4,57
		71	Drug tolerance	0	ja	4,57
		72	Side effects of other medications (MS and other indications)	0	ja	4,57

Tabelle 2: Zuordnung der identifizierten potenziellen Confounder-Variablen zu den jeweiligen potenziellen Confoundern (mehreseitige Tabelle)

Nr. ^a	potenzieller Confounder ^b	Nr. ^c	identifizierte potenzielle Confounder-Variablen ^b	n (Lit ^d)	Int ^e	Score ^f
18	Previous treatment with other disease modifying therapy (DMT)	73	Previous treatment with other disease modifying therapy (DMT)	4	nein	3,86
		74	Type of previous therapy	8	ja	4,29
		75	Treatment naive; treatment history	10	ja	4,14
		76	Duration of previous treatment therapy	1	nein	3,86
		77	Time on first line treatment	1	nein	3,86
		78	Duration of previous natalizumab use	2	nein	4
		79	Number of previous therapies	6	nein	4,29
		80	Past dalfampridine use	1	nein	1,57
		81	Previous treatment with immunosuppressive therapy	1	nein	3,86
		82	Previous treatment with interferon	3	nein	2,71
		83	Time after natalizumab discontinuation	2	nein	4,67
19	patient wish	84	patient wish	0	ja	4,71
		85	Drug route of administration	0	ja	3,86
		86	Willingness to inject a medication permanently	0	ja	4,71
		87	Willingness to take pills permanently	0	ja	4,71
		88	Tolerance of subcutaneous injections	0	ja	4,71
		89	Patient concerns about drug side effects	0	ja	4,29
		90	Patient need for safety	0	ja	4,29

Tabelle 2: Zuordnung der identifizierten potenziellen Confounder-Variablen zu den jeweiligen potenziellen Confoundern (mehreseitige Tabelle)

Nr. ^a	potenzieller Confounder ^b	Nr. ^c	identifizierte potenzielle Confounder-Variablen ^b	n (Lit ^d)	Int ^e	Score ^f
20	Quality of Life	91	EuroQol-5D (EQ-5D)	1	nein	na
		92	MS Impact Scale (MSIS) physical subscale score	1	nein	na
		93	MS Impact Scale (MSIS) psychological subscale score	1	nein	2,14
		94	MS International Questionnaire of Quality of Life (MusiQol)	1	nein	na
		95	Sexual Satisfaction Scale (SSS) score	1	nein	2,43
		96	SF-36 bodily pain subscale score	1	nein	3,57
		97	SF-36 general health subscale score	1	nein	na
		98	SF-36 mental health subscale score	1	nein	na
		99	SF-36 physical functioning subscale score	1	nein	na
		100	SF-36 role emotional subscale score	1	nein	na
		101	SF-36 role physical subscale score	1	nein	na
		102	SF-36 social functioning subscale score	1	nein	na
		103	SF-36 vitality subscale score	1	nein	na
21	Number of previous relapses	104	Number of previous relapses (no defined index time)	6	ja	4,67
		105	Disease severity	1	nein	na
		106	Disease severity at diagnosis	0	ja	4,38
		107	Prognosis	0	ja	4,57
		108	Relapse in prior year	44	nein	4,57
		109	Relapses in prior 2 years	23	nein	4
		110	Relapses in prior 3 years	2	nein	4,14
		111	Relapses in the past 3 months	2	nein	4,43
112	Relapses in the past 6 months	1	nein	4,43		
22	severity of relapses	113	severity of relapses	0	ja	4,71
		114	Past steroid use	1	nein	1,57
		115	Number of treated relapses	1	ja	3,29
		116	Steroid-treated relapses in prior two years	1	nein	3,43
		117	Steroid-treated relapses in prior year	1	nein	3,83

Tabelle 2: Zuordnung der identifizierten potenziellen Confounder-Variablen zu den jeweiligen potenziellen Confoundern (mehreseitige Tabelle)

Nr. ^a	potenzieller Confounder ^b	Nr. ^c	identifizierte potenzielle Confounder-Variablen ^b	n (Lit ^d)	Int ^e	Score ^f
23	Time since diagnosis of MS	118	Time since diagnosis of MS; disease duration	53	nein	3,43
		119	Age at onset of first MS symptoms	1	nein	3,71
		120	Age at start of disease modifying therapy (DMT); Age at start of treatment	2	nein	na
		121	Childhood MS diagnosis	1	nein	3,83
		122	Time between symptom onset and disease modifying therapy (DMT) start	2	nein	3,29
		123	Time since first MS symptoms	11	nein	3,43
		124	Year of treatment initiation	2	nein	3
		125	Age at disease onset	5	nein	4,14
24	Cerebrospinal fluid status	126	Cerebrospinal fluid status	1	ja	2,57
		127	Immunoglobulin (IgM) in cerebrospinal fluid	0	ja	2,43
		128	Neurofilament light chain (NFL)	0	ja	2,86
25	Type of professional work activity	129	Type of professional work activity	0	ja	3,57
26	Educational level	130	Educational level	2	nein	1,57
27	Ethnicity or race	131	Ethnicity or race	13	nein	1,86
28	Pregnancy in the past 3 months	132	Pregnancy in the past 3 months	1	nein	na
<p>a. Nummerierung basierend auf den zusammengefassten inhaltlich überlappenden Confoundern.</p> <p>b. Analog zum Bericht der externen Sachverständigen wurde hier auf die Übersetzung der potenziellen Confounder-Variablen verzichtet, um Übersetzungsfehler zu vermeiden.</p> <p>c. durchgehende Nummerierung aller identifizierten potenziellen Confounder-Variablen von 0 bis 160, für die ausgeschlossenen Variablen von 133 bis 160 siehe Tabelle 3.</p> <p>d. Anzahl der Publikationen, in der die jeweilige potenzielle Confounder-Variable identifiziert wurde.</p> <p>e. Nennung durch die klinischen Expertinnen und Experten im qualitativen Interview (ja oder nein). Im Bericht der externen Sachverständigen erfolgte keine quantitative Erfassung.</p> <p>f. Mittelwert des Umfrage-Scores, Bewertung des Einflusses jeder identifizierten potenziellen Confounder-Variable auf die Therapieentscheidung mithilfe einer 5-Punkte-Skala von 1 (keine Berücksichtigung) bis 5 (sehr starke Berücksichtigung).</p> <p>BMI: Body-Mass-Index; EDSS: Expanded Disability Status Scale; Int.: Interview mit klinischen Expertinnen und Experten; Lit.: Literaturrecherche; MRI: Magnetic resonance imaging; MS: multiple Sklerose; n: Anzahl; Nr: Nummer; SF-36: Short Form-36 Health Survey</p>						

Tabelle 3: Identifizierte Variablen sortiert nach dem jeweiligen Ausschlussgrund (mehreseitige Tabelle)

Nr. ^a	identifizierte Variablen ^b	Ausschlussgrund	n (Lit ^c)	Int ^d	Score ^e
Ausschluss basierend auf fehlender Relevanz im deutschen Versorgungskontext					
133	Health insurance status	keine Relevanz im deutschen Versorgungskontext	1	nein	1,14
134	Health plan type	keine Relevanz im deutschen Versorgungskontext	2	nein	1,14
135	Previous hospital care costs	keine Relevanz im deutschen Versorgungskontext	1	nein	1,5
136	Previous outpatient care costs	keine Relevanz im deutschen Versorgungskontext	2	nein	1,75
Ausschluss basierend auf der Studie CONFIRM (anhand der Ein- / Ausschlusskriterien im Studienprotokoll)					
137	McDonald criteria; MS classification	Einschlusskriterium CONFIRM	5	nein	4,57
138	Hepatitis	Ausschlusskriterium CONFIRM	0	ja	2,63
139	Human immunodeficiency virus (HIV)	Ausschlusskriterium CONFIRM	0	ja	3,25
140	Breastfeeding	Ausschlusskriterium CONFIRM	1	ja	3,57
141	Family planning	Ausschlusskriterium CONFIRM	0	ja	4,43
142	Pregnancy status	Ausschlusskriterium CONFIRM	2	ja	4,86
143	Pregnancy wish	Ausschlusskriterium CONFIRM	0	ja	4,43
144	Timeliness of desire to have children	Ausschlusskriterium CONFIRM	0	ja	4,43
145	Interferon beta treatment within past 3 months	Ausschlusskriterium CONFIRM	1	nein	3
146	Previous treatment with glatiramer acetate	Ausschlusskriterium CONFIRM	3	nein	3,71
147	Reason for dimethyl fumarate treatment	Ausschlusskriterium CONFIRM	1	nein	4,43
148	Years of glatiramer acetate treatment before study	Ausschlusskriterium CONFIRM	1	nein	3,29
149	Medical uncertainty with MS diagnosis	Ausschlusskriterium CONFIRM	0	ja	4
150	Type of MS onset (monofocal or multifocal)	Ausschlusskriterium CONFIRM	1	nein	na
Ausschluss basierend auf methodischen Gründen					
151	EDSS scores/improvement at the end of the study treatment/last check-up	kein Baseline-Kriterium	1	nein	na
152	Time since last visit	kein Baseline-Kriterium	1	nein	2,57
153	Time on study	Studienverlaufscharakteristikum	5	nein	na
154	Time on treatment	Studienverlaufscharakteristikum	2	nein	3,86

Tabelle 3: Identifizierte Variablen sortiert nach dem jeweiligen Ausschlussgrund (mehreseitige Tabelle)

Nr. ^a	identifizierte Variablen ^b	Ausschlussgrund	n (Lit ^c)	Int ^d	Score ^e
Ausschluss da Originalpublikation außerhalb der Fragestellung					
155	Disease modifying therapy (DMT) reinitiation postpartum	Original-Publikation außerhalb der Fragestellung	1	nein	4,14
156	Number of visits 1 year postpartum	Original-Publikation außerhalb der Fragestellung	1	nein	2,71
157	Number of visits in pregnancy	Original-Publikation außerhalb der Fragestellung	1	nein	2,57
158	Age at time of conception	Original-Publikation außerhalb der Fragestellung	1	nein	3,43
159	Age of offspring at last follow-up	Original-Publikation außerhalb der Fragestellung	1	nein	3,14
160	Duration of therapy	Original-Publikation außerhalb der Fragestellung	2	nein	3,86
<p>a. durchgehende Nummerierung aller identifizierten potenziellen Confounder-Variablen von 0 bis 160, für die potenziellen Confounder-Variablen von 0 bis 132 siehe Tabelle 2.</p> <p>b. Analog zum Bericht der externen Sachverständigen wurde hier auf die Übersetzung der potenziellen Confounder verzichtet, um Übersetzungsfehler zu vermeiden.</p> <p>c. Anzahl der Publikationen, in der die jeweilige Variable identifiziert wurde.</p> <p>d. Nennung durch die klinischen Expertinnen und Experten im qualitativen Interview (ja oder nein). Im Bericht der externen Sachverständigen erfolgte keine quantitative Erfassung.</p> <p>e. Mittelwert des Umfrage-Scores, Bewertung des Einflusses jeder einzelnen Variable auf die Therapieentscheidung mithilfe einer 5-Punkte-Skala von 1 (keine Berücksichtigung) bis 5 (sehr starke Berücksichtigung).</p> <p>EDSS: Expanded Disability Status Scale; Int.: Interview mit klinischen Expertinnen und Experten; Lit.: Literaturrecherche; MS: multiple Sklerose; n: Anzahl; Nr.: Nummer</p>					

Anhang C Handlungsempfehlungen – Checkliste

Schritt 1: Identifikation

- Bei der systematischen Informationsbeschaffung sollte zuerst auf Primärpublikationen zu RCTs und Kohortenstudien zurückgegriffen werden. Nachfolgend können auch Leitlinien oder 1-armige Studien berücksichtigt werden.
- Die Ein- und Ausschlusskriterien (PICO) für den Studienpool zur Confounderextraktion sollten sich nah an der zu untersuchenden Fragestellung orientieren.
- Die Extraktion sollte bei vergleichenden Beobachtungsstudien alle Adjustierungsvariablen – unabhängig von der statistischen Signifikanz – umfassen. Bei randomisierten oder nicht vergleichenden Studien sollten alle Patientencharakteristika extrahiert werden.
- Bei der Extraktion aus Primärpublikationen wird die Anwendung eines geeigneten Sättigungskriteriums empfohlen.
- Es sollen Interviews mit klinischen Expertinnen und Experten durchgeführt werden.
- Die klinischen Expertinnen und Experten sollten namentlich benannt und potenzielle Interessenkonflikte offengelegt werden.

Schritt 2: Zusammenführung

- Inhaltlich überlappende Confounder-Variablen sollten zusammengeführt werden. Dies sollte nach der Extraktion erfolgen. Eine transparente und nachvollziehbare Dokumentation sollte gewährleistet sein.
- Für jede identifizierte Confounder-Variable sollte geprüft werden, ob sie innerhalb der zu untersuchenden Fragestellung liegt.
- Der Ausschluss eines potenziellen Confounders aus inhaltlichen Überlegungen muss literaturgestützt begründet werden.
- Für alle identifizierten potenziellen Confounder ist eine Operationalisierung festzulegen.
- Nicht messbare bzw. nicht operationalisierbare potenzielle Confounder können zwar nicht in die Auswertung einbezogen werden, sollten jedoch benannt und bei der Ergebnisinterpretation berücksichtigt werden.

Schritt 3: Relevanzbewertung

- Die Bewertung der potenziellen Confounder durch klinische Expertinnen und Experten sollte als 3. Schritt erfolgen. Eine Bewertung mittels Punktevergabe wird empfohlen.

Erforderliche Dokumentationen

Literaturrecherche

- Angabe der Suchstrategie und der durchsuchten Quellen
- Eindeutige Benennung der Ein- und Ausschlusskriterien
- Angabe eines Flussdiagramms für den Selektionsprozess
- Angabe der Referenzen der als Volltexte geprüften, aber ausgeschlossenen Treffer
- Angabe der Quelle einer jeden extrahierten Confounder-Variable (idealerweise inkl. konkretem Zitat)

Interviews mit klinischen Expertinnen und Experten bzw. Bewertung mittels Umfrage

- Benennung von Namen, Fachgebiet und Institution der eingebundenen Expertinnen und Experten und Offenlegung potenzieller Interessenkonflikte
- Angabe des Interviewleitfadens (oder ähnlichem) und der Unterlagen zur Bewertung der Confounder für die klinischen Expertinnen und Experten

Zusammenführung

- Auflistung der potenziellen Confounder-Variablen, die zu Confoundern zusammengeführt wurden
- bei Ausschluss von identifizierten Confounder-Variablen: literaturgestützte Begründung mit Angabe der Referenz

Veröffentlichung des Berichts

- Die Dokumentation sollte als Teil des Studienprotokolls der geplanten nicht randomisierten vergleichenden Studie veröffentlicht werden.

Anhang D Externe Expertise zum IQWiG-Arbeitspapier

Systematische Confounderidentifikation in der Indikation schubförmig remittierende multiple Sklerose (RRMS)

Externe Expertise zum IQWiG-Arbeitspapier GA23-02

Julian Hirt & Lars G. Hemkens

Pragmatic Evidence Lab
Research Center for Clinical Neuroimmunology and Neuroscience (RC2NB)
Universität Basel und Universitätsspital Basel
Spitalstrasse 2
4031 Basel
Schweiz

Version: 2.0
Stand: 27. November 2024

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	ii
Abbildungsverzeichnis	iii
Abkürzungsverzeichnis.....	iv
1 Hintergrund.....	1
2 Fragestellung.....	2
3 Methoden	3
4 Ergebnisse	11
5 Diskussion	33
6 Fazit	37
7 Literatur	38
Anhang A Suchstrategien	40
Anhang B Interviewleitfaden.....	41
Anhang C Fragebogen	43
Anhang D Einschlussfähige Studienberichte.....	55
Anhang E Ausschlussgründe für Studienberichte	58
Anhang F Datensatz	63

Tabellenverzeichnis

	Seite
Tabelle 1: Auswahlkriterien.....	4
Tabelle 2: Kategorisierung und Klassifizierung potenzieller Confounder	13
Tabelle 3: Nennung von Faktoren im Zusammenhang mit Kausalität.....	18
Tabelle 4: Wichtigkeit potenzieller Confounder	22
Tabelle 5: Cluster potenzieller Confounder	27
Tabelle 6: Geschätzter Zeitaufwand für die Identifizierung potenzieller Confounder	32

Abbildungsverzeichnis

	Seite
Abbildung 1: Literaturrecherche- und Studienaushwahlprozess.....	12

Abkürzungsverzeichnis

Abkürzung	Bedeutung
DAG	Directed Acyclic Graph (Gerichteter azyklischer Graph)
DMF	Dimethylfumarat
DMT	Disease-modifying treatment (Krankheitsmodifizierende Behandlung)
GA	Glatirameracetat
IMRaD	Introduction, Methods, Results, and Discussion
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IQR	Interquartilsbereich
MeSH	Medical Subject Heading
MS	Multiple Sklerose
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RC2NB	Research Center for Clinical Neuroimmunology and Neuroscience Basel
RCT	Randomisierte kontrollierte Studie
RRMS	Schubförmig remittierende multiple Sklerose

Anmerkungen: Abkürzungen, die nicht im Fliesstext und lediglich in den Tabellen verwendet werden, sind nicht im Abkürzungsverzeichnis aufgeführt, sondern unterhalb der entsprechenden Tabelle.

1 Hintergrund

Randomisierte Studien sind der Goldstandard für die Bewertung von Behandlungseffekten [1]. Es gibt jedoch Situationen, in denen eine randomisierte Studie nicht verfügbar ist [2]. Um dennoch Behandlungseffekte zu schätzen, kann dann eine hypothetische Studie entworfen werden (das so genannte "Target Trial") und auf dieser Grundlage eine Beobachtungsstudie durchgeführt werden, die dieses Target Trial simuliert [3]. Dieser sogenannte "Target Trial Emulation" Ansatz umfaßt die Anwendung und Übertragung von Designprinzipien randomisierter kontrollierter Studien (RCT), wobei häufig Routinedaten als primäre Datenquelle verwendet werden [4, 5].

Confounding ist ein großes Problem und ein Risiko für Verzerrungen in solchen nicht-randomisierten Studien, da eine Randomisierung fehlt. Confounding Bias kann definiert werden als Verzerrung des Zusammenhangs zwischen einer Exposition und einem Outcome, weil ein Faktor sowohl mit der Exposition als auch dem Outcome assoziiert ist [3]. Confounding Bias ist ein kritisches Problem für die Validität aller Studien, die versuchen, kausale Effekte zu schätzen, aber nicht randomisiert sind.

Es wird empfohlen, potenzielle Confounder unabhängig, systematisch und umfassend auf der Grundlage von Studien und Expertenmeinungen zu identifizieren [4]. Es wurden verschiedene Methoden zur Identifizierung und Beschreibung bekannter potenzieller Confounder und ihrer Beziehungen vorgeschlagen, wie z.B. gerichtete azyklische Graphen (DAGs) [8], literaturbasierte [5] oder datengesteuerte statistische Ansätze [6]. Es gibt jedoch keinen Konsens über eine standardisierte und systematische Methode zur Identifizierung relevanter Confounder. Im Jahr 2022 haben Pufulete et al. [7] einen mehrstufigen und multimethodischen Ansatz zur Identifizierung von Confoundern entwickelt, der auf einer systematischen Analyse der veröffentlichten Literatur basiert und durch Expertenmeinungen ergänzt wird, die mit empirischen qualitativen und quantitativen Methoden gesammelt wurden.

2 Fragestellung

Unser Ziel war es, systematisch und unabhängig bekannte potenzielle Confounder zu identifizieren, die bei der Bewertung der Auswirkungen krankheitsmodifizierender Behandlungen (DMT) auf die jährliche Schubrate ("relapse rate"), das Fortschreiten der Behinderung ("disability"), die gesundheitsbezogene Lebensqualität ("quality of life") und unerwünschte Ereignisse ("adverse events") bei Erwachsenen mit schubförmig remittierender Multipler Sklerose (RRMS) vorab spezifiziert, bewertet und berücksichtigt werden können. Wir haben uns speziell auf zwei Medikamente konzentriert, Dimethylfumarat (DMF) und Glatirameracetat (GA), die für das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) von besonderem Interesse sind. Die sich daraus ergebende Liste bekannter potenzieller Confounder soll bei der Konzeption und dem Design einer nicht-randomisierten Studie zum Vergleich der Auswirkungen von DMF und GA bei Erwachsenen mit RRMS berücksichtigt werden (die Planung und Durchführung der Studie war nicht Teil dieses Projekts).

3 Methoden

Wir haben eine mehrstufige Studie mit mehreren Methoden durchgeführt, die dem sequenziellen Ansatz von Pufulete et al. [7] folgt. Um bekannte potenzielle Confounder zu identifizieren, haben wir erstens eine systematische Literaturanalyse, zweitens qualitative Interviews mit Klinikern und drittens eine quantitative Umfrage unter Klinikern durchgeführt. In einem vierten Schritt wurden alle identifizierten potenziellen Confounder thematisch kategorisiert und klinisch klassifiziert als (i) Prädiktor für die Intervention und/oder das Outcome oder keines davon und (ii) die erwartete Richtung des Zusammenhangs mit den folgenden Outcomes: jährliche Schubrate, Fortschreiten der Behinderung (gemessen mit der Expanded Disability Status Scale; EDSS), gesundheitsbezogene Lebensqualität und unerwünschte Ereignisse (insbesondere schwerwiegende unerwünschte Ereignisse).

Da wir Daten auf der Grundlage von veröffentlichten Informationen und Interviews oder Umfragen mit Klinikern sammelten und analysierten, fiel unser Projekt nicht unter die Bestimmungen des Schweizer Humanforschungsgesetzes [8, 9]. Wir haben daher keine ethische Bewilligung nach Schweizer Recht beantragt. Fragen im Zusammenhang mit dem Datenschutz und dem Schutz der Privatsphäre bei den Interviews und der Umfrage wurden im Folgenden erläutert. Die Kliniker erhielten keine Vergütung für ihre Teilnahme und hätten jederzeit aus den Interviews und der Umfrage aussteigen können.

Teil I: Literaturanalyse

Auswahlkriterien

Wir schlossen alle Studien und ihre Berichte ein, die alle Einschlusskriterien und keines der in Tabelle 1 aufgeführten Ausschlusskriterien erfüllten. Insgesamt berücksichtigten wir veröffentlichte Berichte über Hauptstudienanalysen, in denen mindestens zwei Interventionen, die DMF und/oder GA beinhalten, in Bezug auf die jährliche Schubrate, das Fortschreiten der Behinderung, die gesundheitsbezogene Lebensqualität und/oder unerwünschte Ereignisse bei Erwachsenen mit RRMS verglichen wurden. Wir haben randomisierte und nicht-randomisierte kontrollierte Studien und Kohortenstudien (unabhängig davon, ob "prospektiv", "retrospektiv", mit gleichzeitigen oder nicht gleichzeitigen Vergleichsgruppen) gemäß der Taxonomie von Viswanathan [10] eingeschlossen. Wir haben veröffentlichte, im Peer-review begutachtete biomedizinische Zeitschriftenartikel mit einer IMRaD-Struktur (Introduction, Methods, Results, and Discussion) in englischer, französischer oder deutscher Sprache ohne Einschränkung des Veröffentlichungsjahres einbezogen.

Tabelle 1: Auswahlkriterien

	Einschlusskriterien	Ausschlusskriterien
Analysis	Main or primary analyses (as such or similarly labelled by authors) with confounder consideration (e.g., by adjustment, stratification, propensity score matching)	Subgroup analyses Sensitivity analyses Economic evaluations Other analyses
Population	Adults (≥18 years) with relapsing-remitting Multiple Sclerosis; at least 90% of all study participants	-
Intervention/exposure	Dimethyl fumarate Glatiramer acetate	-
Comparison	Any	-
Outcome	Annual relapse rate Disability progression (as labelled by authors) Health-related quality of life (as labelled by authors) Adverse events (as labelled by authors)	Time to first relapse Time to disability progression Outcomes labelled as safety or tolerability
Study design	Randomized controlled trials Non-randomized controlled trials Cohort studies (prospective/ retrospective, concurrent/non-concurrent)	Single-arm intervention studies Cross-sectional studies Case reports/series
Publication year	Any	-
Publication type	Peer-reviewed journal article	Preprint Other type (i.e., internet report, conference proceeding, book, thesis, editorial, commentary, letter, study protocol, thesis, registry entry)
Publication structure	Introduction-Methods-Results-and-Discussion (IMRaD)	Abstract-only publication
Publication language	English French German	-

Informationsquellen und Suchstrategie

Wir durchsuchten MEDLINE/PubMed und Embase/Elsevier. Unsere Suchstrategie umfasste synonyme Begriffe und Schlagwörter für die Population (MS) in Kombination mit der Intervention (DMF und GA; Anhang A). Die datenbankspezifischen Suchstrategien wurden von einer Person mit Erfahrung in der Literaturrecherche auf der Grundlage eines bestehenden IQWiG-Berichts zu diesem Thema entwickelt [11] unter Zuhilfenahme des Medical Subject Heading (MeSH) Browsers [12], PubMed PubReMiner [13], designspezifische Suchfilter [14–16] und Polyglot Search [17]. Die endgültigen Suchstrategien wurden von zwei erfahrenen Informationsspezialistinnen aus der Abteilung Informationsmanagement des IQWiG bestätigt. Eine von ihnen führte auch die Recherchen durch und deduplierte die Treffer aus den beiden Datenbanken.

Studienauswahl

Eine Person bewertete die Titel, Zusammenfassungen und Volltexte der Referenzen auf ihre Eignung; im Falle von Unklarheiten wurde diese von einer zweiten Person unterstützt. Die Verfahren zur Studienauswahl wurden in einem Flussdiagramm unter Berücksichtigung der

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) dargestellt [18].

Studienbewertung

Wir haben die Qualität und/oder das Risiko einer Verzerrung der identifizierten Studien nicht kritisch bewertet, da dies für unsere Analyse nicht zielführend wäre [11].

Datenextraktion und -analyse

Eine in der Durchführung systematischer Übersichtsarbeiten erfahrene Person extrahierte die Merkmale der Studie (d.h. das Jahr der Veröffentlichung, das Land, in dem die Studie durchgeführt wurde, und das Studiendesign), die Merkmale der Population zu Beginn der Studie (d.h. Alter und Geschlecht; bei Studien, die eine Matching-Methode verwendeten (z.B. Propensity Score Matching), wurden die Charakteristika der gematchten Kohorten verwendet, sofern verfügbar), Interventionen und alle Faktoren, die zur Adjustierung in den statistischen Haupt- oder Primäranalysen verwendet wurden (unabhängig davon, ob sie statistisch signifikant waren oder nicht). Bei randomisierten Studien identifizierten wir potenzielle Confounder anhand der in tabellarischer Form berichteten Stichprobencharakteristika (siehe "Tabelle 1") [11]. Der Grad der Operationalisierung der identifizierten potenziellen Confounder entsprach dabei den Angaben in der jeweiligen Studie. Eine zweite Person überprüfte und verifizierte alle Extraktionen. Wir vermerkten für alle identifizierten potenziellen Confounder, für welches Outcome sie in der Hauptanalyse verwendet wurden. Wir stützten uns auf den veröffentlichten Studienbericht und die Anhänge, sofern verfügbar. Wenn die Autoren ausdrücklich andere Faktoren im Zusammenhang mit einem kausalen Modell (z.B. Mediatoren) erwähnten und/oder DAGs berichteten, wurden diese Faktoren speziell erfasst und zusammengefasst. Wir haben daher die Volltexte nach "mediator", "modifier", "causal", "mediate", "mediating", "acyclic" und "dag" durchsucht.

Für die Datenextraktion entwickelten wir einen Extraktionsplan, welcher von einer Person anhand einer Stichprobe von 10 zufällig ausgewählten Studien (die in die Studie aufgenommen wurden) getestet wurde. Potenzielle Confounder wurden von zwei Personen mit methodischem und/oder klinischem Fachwissen in Kategorien eingeteilt, die sich aus den Daten ergaben. Die Studien wurden nach Publikationsjahren in zwei Blöcke eingeteilt (veröffentlicht seit 2020 oder davor). Alle in Frage kommenden Studien, die seit 2020 veröffentlicht wurden, wurden extrahiert. Studien, die vor 2020 veröffentlicht wurden, wurden zufällig ausgewählt und auf ihre Eignung hin überprüft und, falls sie geeignet waren, extrahiert. Die Datenextraktion wurde als gesättigt betrachtet, wenn bei der Extraktion von

15 konsekutiven Studien aus der zufällig ausgewählten Stichprobe keine zusätzlichen potenziellen Confounder mehr identifiziert wurden.

Teil II: Qualitative Interviews

Teilnehmer und Rekrutierung

Wir haben klinische Neurologen aus Deutschland und der Schweiz zu individuellen, persönlichen oder online geführten, halbstrukturierten Interviews eingeladen [7]. Mitglieder des Forschungsteams wurden nicht als Interviewteilnehmer berücksichtigt. Wir zogen Kliniker mit Erfahrung in der Verordnung der medikamentösen Behandlungen (DMF oder GA) für Personen mit MS in Betracht. Um potenzielle Interviewteilnehmer zu identifizieren, stützten wir uns auf das große nationale und internationale Netzwerk von MS-Experten an unserer Einrichtung, dem Research Center for Clinical Neuroimmunology and Neuroscience Basel (RC2NB; Universitätsspital Basel und Universität Basel). Insbesondere haben wir die lokalen Studienleiter der Schweizer MS-Kohortenstudie eingeladen, an der die 8 größten MS-Zentren der Schweiz beteiligt sind [19], sowie die klinischen Neurologen, die Projektpartner unserer landesweiten pragmatischen Plattformstudie "MultiSCRIPT" sind [20] sowie ausgewählte Kontakte über die beruflichen Netzwerke der klinischen MS-Experten des RC2NB.

Potenzielle Interviewteilnehmer wurden per E-Mail, Telefon oder persönlich eingeladen. Kliniker, die ihr Interesse bekundeten, erhielten per E-Mail ein Informationsblatt zur Studie. Sobald ein Interviewtermin vereinbart war, erhielten die Kliniker schriftliche Studieninformationen (d.h. Ziel und Methoden der Studie, Datenverarbeitung, Datenschutz, Datensicherheit und Datenspeicherung) und unterschrieben vor der Teilnahme (persönlich oder elektronisch) eine Einverständniserklärung. Die Teilnahme an der Studie war nicht verpflichtend, und konnte jederzeit auch nach Zustimmung zur Teilnahme widerrufen werden.

Datenerhebung

Wir entwickelten einen vignettenbasierten Interviewleitfaden mit einem MS-erfahrenen klinischen Wissenschaftler aus unserem Team. Dieser Leitfaden wurde in einem iterativen Feedback-Prozess von erfahrenen Neurologen und MS-Experten überprüft, die Forschungsgruppenleiter an der Universität Basel und Teil des RC2NB waren. Der auf Vignetten basierende Interviewleitfaden umfasste (i) eine offene Einleitungsfrage, die nach dem ersten Interview hinzugefügt wurde, um die Interviews angemessen zu beginnen und die Befragten in das Gespräch einzubeziehen, und (ii) 5 Szenarien zu klinischen Fällen und Indikationen, welche die Diskussionen mit den Klinikern über die Verordnungsentscheidungen

und Überlegungen zu DMF und GA initiieren sollten (Anhang B). Die Kliniker wurden gefragt, ob sie DMF und GA für jedes bestimmte Szenario verschreiben würden, wie sie das Medikamentenregime wählen und welche Faktoren ihre Verordnung beeinflussen. Wir haben speziell offene Fragen zu krankheitsbezogenen Faktoren und zu patientenbezogenen Faktoren gestellt. Darüber hinaus erfassten wir Merkmale der Teilnehmer und des klinischen Kontextes (Jahre sowie aktueller Standort und Setting, geschätzte Anzahl der Verordnungen von DMF und GA im letzten Jahr und Verwendung von Leitlinien oder lokalen Standards für die Verordnung).

Alle Interviews wurden aufgezeichnet; die Aufnahmen wurden von einem professionellen Transkriptionsdienst (Transkripto) wortwörtlich transkribiert.

Datenanalyse

Alle Aufnahmen wurden abgehört und gleichzeitig wurden die entsprechenden Transkripte von einer Person, d.h. der gleichen Person, die das Interview geführt hat, auf ihre Richtigkeit überprüft. Die anonymisierten Transkripte wurden in MAXQDA (Version 2022 [21]) zur Datenanalyse importiert. Eine in der qualitativen Forschung und Datenanalyse erfahrene Person folgte einem schrittweisen Ablauf für die qualitative Datenanalyse mit dem Ziel, eine Liste von Faktoren zu identifizieren, die Kliniker bei der Verordnung von DMF und GA beeinflussen. Der Analyserahmen umfasste die zeilenweise Kodierung (Paraphrasierung oder Etikettierung) von wichtigen Textinformationen (Codes), die Kategorisierung der Codes und die Darstellung der Kategorien und Codes in einer Rahmenmatrix [26]; ein Ansatz, der bereits in der Primär- und Sekundärforschung angewandt wurde [27–29]. Die identifizierten potenziellen Confounder wurden dann von zwei Personen mit methodischem und/oder klinischem Fachwissen mit den bei der Literaturanalyse (Teil I) identifizierten potenziellen Confounder abgeglichen und kategorisiert.

Teil III: Quantitativer Survey

Teilnehmer und Rekrutierung

Wir haben Kliniker aus Deutschland und der Schweiz für die Teilnahme an der Umfrage rekrutiert. Wir schickten die Einladung zur Umfrage an (i) Kliniker, die wir über Websites und die Literatur identifiziert hatten (d.h. deutsche und schweizerische MS-Zentren) und (ii) klinische MS-Experten des RC2NB mit der Bitte, die Einladung an ihre beruflichen Netzwerke weiterzuleiten. Wir schickten nach einer Woche eine Erinnerung an jeden der

Einladungsempfänger. Mitglieder des Forschungsteams und Interviewpartner aus Teil II wurden nicht als Umfrageteilnehmer berücksichtigt.

Die Webseite der Umfrage enthielt Informationen über Ziel und Methoden der Studie, Datenverarbeitung, Schutz, Sicherheit und Speicherung. Mit der Teilnahme an der Umfrage erklärten sich die Teilnehmer auch mit diesen Aspekten einverstanden. Unsere Rekrutierungsstrategie ließ prinzipiell zu (auch wenn dies nicht beabsichtigt war), dass Interviewpartner aus Teil II eine Einladung zur Umfrage erhielten. Daher enthielt die Landing Page der Website eine Eingangsfrage zur Teilnahme an den qualitativen Interviews, welche ggf. direkt zum Ende der Umfrage geleitet hat. Damit wollten wir vermeiden, dass ein Teilnehmer sowohl an den qualitativen Interviews als auch an der quantitativen Umfrage teilnimmt.

Datenerhebung und -analyse

Wir entwickelten einen Fragebogen, in dem alle bekannten potenziellen Confounder, die durch die Literaturanalyse und die qualitativen Interviews identifiziert wurden, alphabetisch nach Confounder-Kategorie und innerhalb jeder Kategorie nach potenziellem Confounder sortiert sind. Die Teilnehmer wurden gebeten, die Bedeutung jedes potenziellen Confounders zu bewerten, d.h. wie stark jeder einzelne Faktor ihre Entscheidung bei der Verordnung von DMF oder GA auf einer 5-Punkte-Skala von '1 - keine Berücksichtigung' bis '5 - sehr starke Berücksichtigung' beeinflusst (Anhang C). Aufgrund der großen Anzahl einzelner potenzieller Confounder, die in der Literatur und den qualitativen Interviews identifiziert wurden, und um die Zeit, die die Teilnehmer für das Ausfüllen der Umfrage benötigen, zu minimieren, haben wir ähnliche potenzielle Confounder für den Survey zusammengefasst (z.B. Messinstrument-Subskalen; n=13) oder sie weggelassen, wenn sie ohne einen spezifischen Studienkontext nicht ausreichend verständlich waren (z.B. Studiendauer; n=2). Zudem lagen wenige potenzielle Confounder zum Start des Surveys noch nicht vor, da die Teilschritte des Projektes teilweise parallel durchgeführt wurden (n=8; siehe Abweichungen vom Protokoll). Die Teilnehmer konnten über ein Freitextfeld alle potenziellen Confounder hinzufügen, die nicht bereits aufgeführt waren. Zudem wurden die beruflichen Merkmale der Teilnehmer (Berufserfahrung in Jahren [allgemein und MS-spezifisch] sowie der aktuelle Standort und das Setting) erfasst. Zwei Mitglieder des Forschungsteams und MS-Experten des RC2NB testeten die Umfrage vorab, um die Benutzerfreundlichkeit und die inhaltliche Validität zu beurteilen.

Wir haben Unipark/EFS Survey verwendet, um die Umfrage zu entwerfen und sie per E-Mail zu versenden [22]. Nach zehn Tagen wurden Erinnerungen an die Adressen geschickt, die wir ursprünglich kontaktiert hatten. Die Antworten haben wir deskriptiv ausgewertet.

Teil IV: Kategorisierung und Klassifizierung

Alle identifizierten potenziellen Confounder wurden auf der Grundlage klinischer Überlegungen in eine der vier folgenden Kategorien eingeteilt: Prädiktor für Intervention und Outcome, Prädiktor nur für die Intervention, Prädiktor nur für das Outcome und weder Prädiktor für Intervention noch für das Outcome. Wir gaben auch die erwartete Richtung des Zusammenhangs mit den folgenden Outcomes an: jährliche Schubrate, Fortschreiten der Behinderung (gemessen am EDSS), gesundheitsbezogene Lebensqualität und unerwünschte Ereignisse (insbesondere schwerwiegende unerwünschte Ereignisse). Wir fügten der Richtung des Zusammenhangs der endgültigen Liste der potenziellen Confounder eine Bewertung der Einschätzungssicherheit hinzu.

Die Klassifizierungsurteile basierten auf der klinischen und methodischen Expertise des Teams und in Zusammenarbeit mit zwei klinischen MS-Experten (Senior Consultants Neurologie) des RC2NB. Die Ergebnisse wurden in tabellarischer Form dargestellt [7].

Um die endgültige Liste der potenziellen Confounder so übersichtlich wie möglich zu gestalten, haben wir individuelle potenzielle Confounder, die auf ähnliche Weise gemessen oder bewertet werden können, in Cluster zusammengefasst.

Machbarkeitsanalyse

Wir schätzten die personellen Ressourcen, die zur Erfüllung der jeweiligen Aufgaben erforderlich waren (d.h. identifizierte potenzielle Confounder und Zeit [Personentage] pro angewandte Methode; Teil I-III).

Abweichungen vom Protokoll

Es gab keine wesentlichen Abweichungen vom Protokoll, das vom IQWiG geprüft und akzeptiert wurde. Für die Literaturübersicht haben wir die Ausschlusskriterien in Bezug auf die Ergebnisse genauer spezifiziert (d.h. wir haben die Zeit bis zum Rückfall [time to relapse] und die Zeit zum Fortschreiten der Behinderung nicht einbezogen [time to disease progression]), und wir haben die in Frage kommenden Publikationssprachen aufgrund der Sprachkenntnisse des endgültigen Forschungsteams auf Englisch, Deutsch und Französisch beschränkt. Bei den qualitativen Interviews wurde nach dem ersten Gespräch eine offene Einleitungsfrage hinzugefügt, um die Interviews angemessen zu beginnen und die Befragten in das Gespräch

einzu beziehen. Wir entschieden uns für acht und nicht wie geplant für zehn Teilnehmer, da die Rekrutierung schwierig war und wir feststellten, dass keine wesentlichen neuen Erkenntnisse gewonnen wurden, die Beiträge dieser Experten also "gesättigt" waren, so dass die gewonnenen Ergebnisse die Faktoren, die von Klinikern bei der Verordnung von DMF oder GA berücksichtigt werden, ausreichend widerspiegeln. Im Projektverlauf wurden zeitliche Überlappungen bei der Ausführung der einzelnen methodischen Schritte (Literaturanalyse, qualitative Interviews, quantitativer Survey) notwendig, was dazu führte, dass wir den quantitativen Survey bereits initiierten bevor der Abgleich der aus der Literatur identifizierten potenziellen Confounder durch eine zweite Person sowie die Klassifizierung und Kategorisierung der bereits identifizierten potenziellen Confounder abgeschlossen war. Dadurch waren wenige aus der Literatur und den qualitativen Interviews identifizierten potenziellen Confounder nicht in ihrer finalen Benennung im quantitativen Survey berücksichtigt, wir haben es jedoch für wichtig erachtet, diese im Endergebnis zu berücksichtigen. Für den qualitativen Survey wollten wir 30 Personen rekrutieren, was trotz erheblicher Zeitaufwendung und Anstrengungen zur Rekrutierung ehrenamtlicher Experten in dem vorgegebenen Zeitrahmen nicht möglich war. Ein Hauptgrund dafür war, dass wir die Teilnehmer bitten, ihre Einschätzung zu 137 einzelnen potenziellen Confoundern abzugeben, was sehr zeitaufwendig ist. Außerdem fiel die Umfrage in die Sommer- und Hauptferienzeit. Um die Durchführbarkeit der Umfrage zu verbessern und weitere Hürden bei der Rekrutierung zu vermeiden, haben wir das Ranking der zehn wichtigsten Confounder ausgelassen und die Teilnehmer stattdessen aufgefordert, die Wichtigkeit für jeden Confounder einzeln einzuschätzen.

4 Ergebnisse

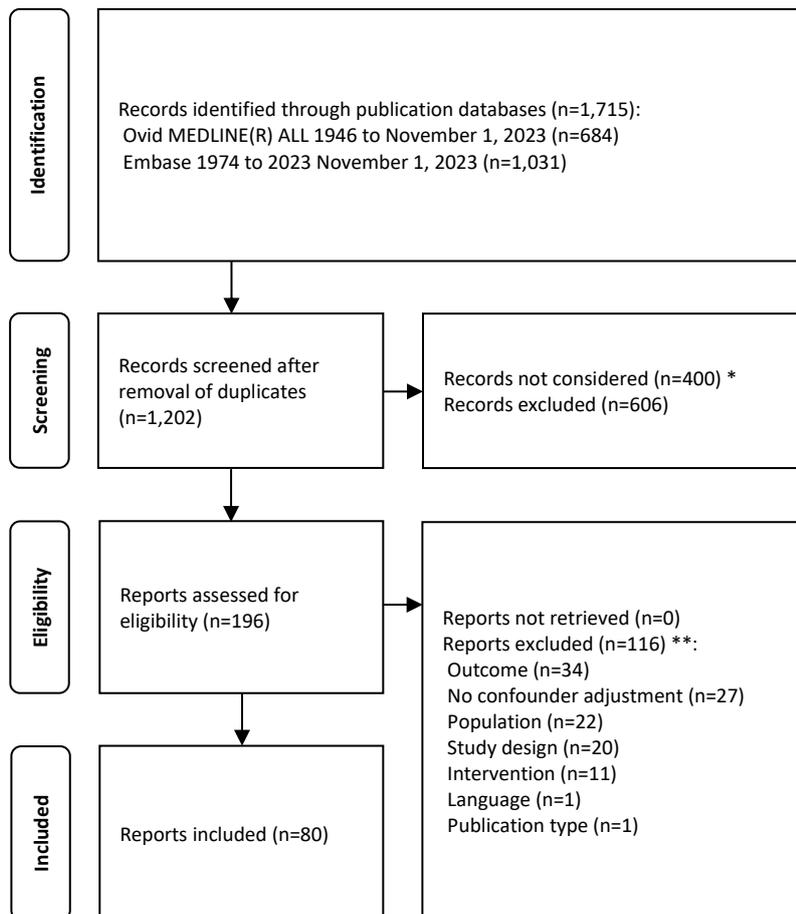
Im Folgenden stellen wir die detaillierten Ergebnisse der einzelnen Studienteile vor.

Teil I: Literaturanalyse

Die systematische Literatursuche ergab insgesamt 1.715 Treffer, die nach der Deduplizierung 1.202 einzelne Referenzen ergaben. Davon wurden 351 im Jahr 2020 oder später veröffentlicht (Sammlung "2020-2023") und 851 vor 2020 (Sammlung "pre2020"). Aus der Sammlung "2020-2023" haben wir alle 351 Titel und Zusammenfassungen gesichtet und 110 Volltexte als potenziell geeignet eingestuft, von denen wir 40 nach der Bewertung der Volltexte eingeschlossen haben. Aus der "pre2020"-Sammlung haben wir 451 der 851 Titel und Zusammenfassungen und 86 Volltexte nach dem Zufallsprinzip gescreent, um unser Sättigungskriterium zu erreichen (d.h. kein zusätzlicher potenzieller Confounder wurde in 15 aufeinanderfolgenden einschlussfähigen Studien identifiziert), nachdem wir die Daten aus 40 Studien extrahiert hatten (Abbildung 1). Die einschlussfähigen Studienberichte sind in Anhang D aufgeführt. Die Gründe für den Ausschluss einzelner Volltextberichte sind in Anhang E aufgeführt.

Wir identifizierten und kategorisierten 125 potenzielle Confounder aus der Literatur (Tabelle 2); 57 (46%) stammten nur aus Studien, die 2020 oder später veröffentlicht wurden, 26 (21%) nur aus Studien vor 2020 und 42 (34%) aus Studien vor und nach 2020. 83 (66%) wurden nur in Beobachtungsstudien berichtet, 13 (10%) nur in randomisierten Studien und 29 (23%) in beiden Studiendesigns.

Konzepte kausaler Modellierung wurden selten berichtet, so fanden wir den Begriff "causal" nur in 3 von 80 in Frage kommenden Volltexten und den Begriff "acyclic" in 1 von 80 in Frage kommenden Volltexten (Tabelle 3). "Causal" wurde verwendet, um Studienergebnisse zu interpretieren oder um kausale Annahmen zu beschreiben. "Mediator", "modifier", "mediate", "mediating" oder "DAG" wurden im Haupttext der einschlussfähigen Studienberichte nie erwähnt.



Anmerkungen: * Da unser Sättigungskriterium erreicht war, haben wir 400 Datensätze aus der "pre2020"-Sammlung nicht berücksichtigt; ** Ausschlussgründe pro Studienbericht sind in Anhang E hinterlegt.

Abbildung 1: Literaturrecherche- und Studienaushwahlprozess

Tabelle 2: Kategorisierung und Klassifizierung potenzieller Confounder

Potenzielle Confounder (n=160), die im Rahmen der Literaturanalyse, der qualitativen Interviews und/oder des quantitativen Surveys identifiziert wurden; alphabetisch und nach Kategorie des Confounders sortiert.¹ Die detaillierte Darstellung der potenziellen Confounder je Datenquelle und ihre Klassifizierung befindet sich in Anhang F (Mappen "Confounder-Extraktion" und "Confounder-Klassifizierung").

Individuelle Confounder pro Kategorie (und alternative Ausdrücke, falls zutreffend)	n	Quelle	Intervention	Outcome	Prädiktor Intervention	Prädiktor Outcome	Erwartete Richtung der Assoziation	Sicherheit
Clinical status (n=15)								
Ambulation index	3	rand	ga	dis; rr	py	py	neg	exp rat
Clinical condition	1	int	dmf; ga	-	py	py	neg	exp rat
Depression diagnosis	2	obs	ga	rr	py	py	neg	exp rat
Deterioration index (defined as the Expanded Disability Status Scale (EDSS) score divided by the disease duration)	2	rand	ga	dis; ae	py	py	neg	exp rat
Disability status	2	obs; int	dmf; ga	dis; rr	py	py	neg	exp rat
Expanded Disability Status Scale (EDSS) change preceding year	1	obs	dmf	dis; rr	py	py	neg	exp rat
Expanded Disability Status Scale (EDSS) in the past 3 months	1	obs	ga	dis	py	py	neg	exp rat
Expanded Disability Status Scale (EDSS) score (mean or median)	61	obs; rand	dmf; ga	dis; rr; qol; ae	py	py	neg	exp rat
Expanded Disability Status Scale (EDSS) scores/improvement at the end of the study treatment/last check-up	1	obs	ga	dis; rr; qol	py	py	neg	exp rat
Hamilton Anxiety Rating Scale (HAM-A)	1	obs	ga	dis; qol	py	py	neg	exp rat
Hamilton Depression Rating Scale (HAM-D)	1	obs	ga	dis; qol	py	py	neg	exp rat
Multiple Sclerosis Functional Composite (MSFC) score	3	obs; rand	dmf; ga	dis; rr; qol; ae	py	py	neg	exp rat
Multiple Sclerosis Severity Score (MSSS)	1	obs	ga	dis	py	py	neg	exp rat
Neurological symptoms	1	int	dmf; ga	-	py	py	neg	exp rat
Pyramidal impairment	2	obs; int	dmf; ga	dis	py	py	neg	exp rat
Comorbidity (n=14)								
Allergy	1	obs	ga	dis	py	py	neg	exp rat
Cardiovascular comorbidities	2	obs; int	dmf	dis	py	py	neg	exp rat
Charlson Comorbidity Index (CCI) score	3	obs	dmf; ga	rr	py	py	neg	exp rat
Chronic infections	1	int	dmf; ga	-	py	py	neg	exp rat
Comorbidities (general)	3	all	dmf; ga	dis; rr	py	py	neg	exp rat
Dermatological diseases	1	int	dmf	-	py	py	neg	exp rat
Diabetes	1	obs	ga	rr	py	py	neg	exp rat
Headache	1	obs	ga	dis	py	py	neg	exp rat
Hepatitis	1	int	dmf; ga	-	py	py	neg	exp rat
Hospitalisations in previous year	1	obs	ga	rr	py	py	neg	exp rat
Human immunodeficiency virus (HIV)	1	int	dmf; ga	-	py	py	neg	exp rat
Skin problems	1	int	dmf; ga	-	py	py	neg	exp rat

¹ Englisch war die Projektarbeitsprache. Auf die Übersetzung der Confounder und Confounder-Kategorien wurde verzichtet, um Übersetzungsfehler zu vermeiden.

Thyroid dysfunction	1	obs	ga	dis	py	py	neg	exp rat
Tuberculosis	1	int	dmf; ga	-	py	py	neg	exp rat
Costs (n=2)								
Previous hospital care costs	1	obs	dmf	dis; rr	unclear	unclear	unclear	-
Previous outpatient care costs	2	obs	dmf; ga	dis; rr	unclear	unclear	unclear	-
Disease status (clinical and magnetic resonance imaging; n=4)								
Disease activity	2	obs	dmf	dis; rr; ae	py	py	neg	exp rat
Disease course	1	obs	dmf; ga	dis; rr	py	py	neg	exp rat
Disease severity	1	obs	ga	rr	py	py	neg	exp rat
Disease severity at diagnosis	1	int	dmf; ga	-	py	py	neg	exp rat
Family planning, pregnancy, and birth (n=11)								
Age at time of conception	1	obs	ga	ae	py	py	unclear	-
Age of offspring at last follow-up	1	obs	ga	ae	py	py	unclear	-
Breastfeeding	2	obs; int	dmf; ga	dis; rr	py	py	unclear	-
Disease modifying therapy (DMT) reinitiation postpartum	1	obs	dmf; ga	dis; rr	py	py	unclear	-
Family planning	1	int	dmf; ga	-	py	py	unclear	-
Number of visits 1 year postpartum	1	obs	dmf; ga	dis; rr	py	py	unclear	-
Number of visits in pregnancy	1	obs	dmf; ga	dis; rr	py	py	unclear	-
Pregnancy in the past 3 months	1	obs	ga	dis	py	py	unclear	-
Pregnancy status	3	obs; int	dmf; ga	dis; rr	py	py	unclear	-
Pregnancy wish	1	int	dmf; ga	-	py	py	unclear	-
Timeliness of desire to have children	1	int	dmf; ga	-	py	py	unclear	-
Geographical location and study center (n=4)								
Country	6	obs; rand	dmf; ga	dis; rr	py	unclear	unclear	-
Region	5	obs; rand	dmf; ga	dis; rr; ae	py	unclear	unclear	-
Study center; Multiple sclerosis (MS) center	3	obs	dmf; ga	dis; rr	py	unclear	unclear	-
Within-country region	2	obs	dmf; ga	dis; rr	py	unclear	unclear	-
Health insurance type and status (n=2)								
Health insurance status	1	obs	dmf	dis; qol	unclear	unclear	unclear	-
Health plan type	2	obs	ga	rr	unclear	unclear	unclear	-
Lab test (n=4)								
Absolute lymphocyte count; Lymphopenia	2	obs; int	dmf	ae	py	py	neg	exp rat
Cerebrospinal fluid status	2	obs; int	dmf; ga	dis; rr; qol	py	py	unclear	-
Immunoglobulin (IgM) in cerebrospinal fluid	1	int	dmf; ga	-	py	py	neg	exp rat
Neurofilament light chain (NFL)	1	int	dmf; ga	-	py	py	neg	exp rat
Lifestyle (n=5)								
Body mass index (BMI); or height and weight	6	obs; rand	dmf; ga	dis; rr; ae	py	unclear	unclear	-
Overweight or obesity	1	obs	ga	rr	py	unclear	unclear	-
Tobacco use	1	obs	dmf	rr; ae	unclear	unclear	unclear	-
Type of professional work activity	1	int	dmf; ga	-	py	unclear	unclear	-
Weight	2	rand	dmf; ga	rr; ae	py	unclear	unclear	-

Magnetic resonance imaging (n=17)									
Brain volume; normalized brain volume	8	all	dmf; ga	dis; rr; qol; ae	py	py	neg	exp rat	
Enhancing lesion volume	3	rand	ga	dis; rr	py	py	neg	exp rat	
Lesion load	1	int	dmf; ga	-	py	py	neg	exp rat	
Localisation of lesions	1	int	dmf; ga	-	py	py	neg	exp rat	
Magnetic resonance imaging (MRI) brain lesions; number of lesions	3	all	dmf; ga	dis; rr; qol	py	py	neg	exp rat	
Magnetic resonance imaging (MRI) status	2	obs; int	dmf; ga	dis; rr; qol	py	py	unclear	-	
New lesions	1	int	ga	-	py	py	neg	exp rat	
Normalized grey matter volume	1	rand	ga	rr; ae	py	py	neg	exp rat	
Normalized white matter volume	1	rand	ga	rr; ae	py	py	neg	exp rat	
Number of gadolinium enhancing lesions; Number of contrast-enhancing lesions; Number of patients with gadolinium enhancing lesions; Patients with enhancement	22	all	dmf; ga	dis; rr; qol; ae	py	py	neg	exp rat	
Number of T2 lesions	12	obs; rand	dmf; ga	dis; rr; qol; ae	py	py	neg	exp rat	
Presence of at least 9 magnetic resonance imaging (MRI) T2 lesions and at least one contrast-enhancing MRI lesion	2	obs	dmf	dis; rr	py	py	neg	exp rat	
Size of lesions	1	int	dmf; ga	-	py	py	neg	exp rat	
Spinal lesions	1	int	dmf; ga	-	py	py	neg	exp rat	
Status of gadolinium enhancing T1 activity	1	obs	ga	dis; rr	py	py	neg	exp rat	
T1 hypointense lesion volume	7	rand	dmf; ga	dis; rr; qol; ae	py	py	neg	exp rat	
Volume of T2 lesions	7	obs; rand	dmf; ga	dis; rr; qol; ae	py	py	neg	exp rat	
Multiple sclerosis disease duration (n=6)									
Age at onset of first multiple sclerosis symptoms	1	rand	ga	dis; rr; ae	py	py	unclear	-	
Age at start of disease modifying therapy (DMT); Age at start of treatment	2	obs	ga	dis; rr; qol	py	py	unclear	-	
Childhood multiple sclerosis diagnosis	1	rand	dmf	dis; rr; qol; ae	py	py	neg	exp rat	
Time between symptom onset and disease modifying therapy (DMT) start	2	obs	ga	dis; rr; qol	py	py	neg	exp rat	
Time since diagnosis of multiple sclerosis; disease duration	53	obs; rand	dmf; ga	dis; rr; qol; ae	py	py	neg	exp rat	
Time since first multiple sclerosis symptoms	11	obs; rand	dmf; ga	dis; rr; qol; ae	py	py	unclear	-	
Multiple sclerosis therapy (n=23)									
Drug efficacy (known or suspected)	1	int	dmf; ga	-	py	py	unclear	-	
Duration of previous natalizumab use	2	obs	dmf	dis; rr	py	py	unclear	-	
Duration of previous treatment therapy	1	obs	dmf	dis; rr; qol	py	py	unclear	-	
Duration of therapy	2	obs	dmf; ga	dis; rr	py	py	unclear	-	
Interferon beta treatment within past 3 months	1	rand	dmf	dis; rr	py	py	unclear	-	
Number of previous therapies	6	obs	dmf; ga	dis; rr; ae	py	py	unclear	-	
Past dalfampridine use	1	obs	ga	rr	unclear	unclear	unclear	-	
Past steroid use	1	rand	ga	dis; rr	unclear	unclear	unclear	-	
Previous treatment with glatiramer acetate	3	obs; rand	dmf; ga	dis; rr; qol	py	py	unclear	-	
Previous treatment with immunosuppressive therapy	1	obs	dmf	dis; rr; ae	py	py	unclear	-	

Previous treatment with interferon	3	obs; rand	dmf; ga	dis; rr; qol	py	py	unclear	-
Previous treatment with other disease modifying therapy (DMT)	4	obs; rand	dmf; ga	dis; rr	py	py	unclear	-
Reason for dimethyl fumarate treatment	1	obs	dmf	dis; rr	unclear	unclear	unclear	-
Reason for discontinuing prior therapy	2	obs	dmf; ga	dis; rr	py	py	unclear	-
Time after natalizumab discontinuation	2	obs	dmf; ga	dis; rr	py	unclear	unclear	-
Time on first line treatment	1	obs	dmf	dis; rr	py	py	unclear	-
Time on study	5	obs; rand	dmf; ga	dis; rr; qol; ae	unclear	unclear	unclear	-
Time on treatment	2	obs; rand	dmf	dis; rr; ae	py	unclear	unclear	-
Time since last visit	1	obs	dmf	dis; rr	py	unclear	unclear	-
Treatment naive; treatment history	11	all	dmf; ga	dis; rr; qol; ae	py	py	unclear	-
Type of previous therapy	9	all	dmf; ga	dis; rr; qol	py	py	unclear	-
Year of treatment initiation	2	obs	ga	dis; rr	py	unclear	unclear	-
Years of glatiramer acetate treatment before study	1	rand	ga	ae	py	py	unclear	-
Multiple sclerosis type and diagnosis criteria (n=3)								
MacDonald criteria; MS classification	5	obs; rand	dmf; ga	dis; rr	py	py	neg	exp rat
Medical uncertainty with multiple sclerosis diagnosis	1	int	ga	-	py	unclear	unclear	-
Type of multiple sclerosis (MS)	1	obs	dmf; ga	dis	na	na	na	-
Patient preferences, willingness, and concerns (n=6)								
Drug route of administration	1	int	dmf; ga	-	py	unclear	unclear	-
Patient concerns about drug side effects	1	int	dmf	-	py	unclear	unclear	-
Patient need for safety	1	int	ga	-	py	unclear	unclear	-
Patient wish	1	int	dmf; ga	-	py	unclear	unclear	-
Willingness to inject a medication permanently	1	int	ga	-	py	unclear	unclear	-
Willingness to take pills permanently	1	int	dmf; ga	-	py	unclear	unclear	-
Patient-reported symptoms, health, functioning and well-being (n=21)								
Bowel symptoms	1	obs	ga	rr	py	py	neg	exp rat
Cognition (self-rated)	1	obs	dmf	dis	py	py	neg	exp rat
Depression (self-rated)	1	obs	dmf	dis	py	py	neg	exp rat
EuroQol-5D (EQ-5D)	1	obs	ga	dis; rr; qol	py	py	neg	exp rat
Fatigue	1	obs	ga	rr	py	py	neg	exp rat
Functional System (FS) score	1	obs	dmf	dis; rr	py	py	neg	exp rat
Modified Fatigue Impact Scale (MFIS) score	1	obs	ga	dis; qol	py	py	neg	exp rat
Multiple Sclerosis Impact Scale (MSIS) physical subscale score	1	obs	dmf	dis; qol	py	py	neg	exp rat
Multiple Sclerosis Impact Scale (MSIS) psychological subscale score	1	obs	dmf	dis; qol	py	py	neg	exp rat
Multiple Sclerosis International Questionnaire of Quality of Life (MusiQoL)	1	obs	ga	dis; qol	py	py	neg	exp rat
Numbness	1	obs	ga	rr	py	py	neg	exp rat
Patient Determined Disease Steps (PDDS) score	2	obs	dmf	dis; qol	py	py	neg	exp rat
Sexual Satisfaction Scale (SSS) score	1	obs	ga	dis; qol	py	py	neg	exp rat
Short Form-36 (SF-36) bodily pain subscale score	1	obs	ga	dis; qol	py	py	neg	exp rat

Short Form-36 (SF-36) general health subscale score	1	obs	ga	dis; qol	py	py	neg	exp rat
Short Form-36 (SF-36) mental health subscale score	1	obs	ga	dis; qol	py	py	neg	exp rat
Short Form-36 (SF-36) physical functioning subscale score	1	obs	ga	dis; qol	py	py	neg	exp rat
Short Form-36 (SF-36) role emotional subscale score	1	obs	ga	dis; qol	py	py	neg	exp rat
Short Form-36 (SF-36) role physical subscale score	1	obs	ga	dis; qol	py	py	neg	exp rat
Short Form-36 (SF-36) social functioning subscale score	1	obs	ga	dis; qol	py	py	neg	exp rat
Short Form-36 (SF-36) vitality subscale score	1	obs	ga	dis; qol	py	py	neg	exp rat
Prescriber (n=1)								
Type and speciality of prescriber	2	obs	dmf; ga	rr; ae	py	unclear	unclear	-
Prognosis (n=1)								
Prognosis	1	int	ga	-	py	py	neg	exp rat
Relapse (n=12)								
Number of previous relapses (no defined index time)	7	all	dmf; ga	dis; rr; qol	py	py	neg	exp rat
Number of treated relapses	2	obs; int	dmf; ga	dis; rr	py	py	neg	exp rat
Relapse in prior 3 years	2	rand	dmf	dis; rr	py	py	neg	exp rat
Relapses in prior 2 years	23	obs; rand	dmf; ga	dis; rr; qol; ae	py	py	neg	exp rat
Relapses in prior year	44	obs; rand	dmf; ga	dis; rr; qol; ae	py	py	neg	exp rat
Relapses in the past 3 months	2	obs	ga	dis; rr	py	py	neg	exp rat
Relapses in the past 6 months	1	obs	dmf	dis	py	py	neg	exp rat
Remit of relapse	1	int	dmf; ga	-	py	py	neg	exp rat
Severity of relapses	1	int	dmf; ga	-	py	py	neg	exp rat
Steroid-treated relapses in prior two years	1	obs	ga	dis; rr	py	py	neg	exp rat
Steroid-treated relapses in prior year	1	obs	ga	dis; rr	py	py	neg	exp rat
Time since last relapse	5	obs; rand	dmf; ga	dis; rr; ae	py	py	neg	exp rat
Socio-demographic characteristics (n=5)								
Age	71	all	dmf; ga	dis; rr; qol; ae	py	py	unclear	-
Age at disease onset	5	obs; rand	dmf; ga	dis; rr; qol; ae	py	py	unclear	-
Educational level	2	obs	dmf	dis; rr	unclear	unclear	unclear	-
Ethnicity or race	13	obs; rand	dmf; ga	dis; rr; qol; ae	unclear	unclear	unclear	-
Gender/sex	62	all	dmf; ga	dis; rr; qol; ae	py	py	neg (female)	exp rat
Tolerance and side effects (n=4)								
Drug tolerance	1	int	dmf; ga	-	py	py	unclear	-
Gastrointestinal disorders	1	int	dmf	-	py	py	unclear	-
Side effects of other medications (multiple sclerosis and other indications)	1	int	dmf; ga	-	py	py	unclear	-
Tolerance of subcutaneous injections	1	int	ga	-	py	py	unclear	-

Abkürzungen: - = not specified/assessed; all = Potential confounder source is observational study, randomized trial and qualitative interview; dis = disability; DMF = Dimethyl fumarate; exp rat = Expert rating; GA = Glatiramer acetate; int = qualitative interview; n = number; na = not applicable; neg = negative; obs = observational study; rand = randomized trial; rr = relapse rate.

Tabelle 3: Nennung von Faktoren im Zusammenhang mit Kausalität

Causal (n=3)	Zitat	Textstelle	Zusammenfassung
Buron 2021 [23]	"For the design allowing time-varying exposure and a common baseline definition, we applied marginal structural Cox models to deal with time-dependent confounding. ¹³ This approach was chosen as time-varying confounders could be affected by previous decisions on whether to switch treatment, potentially rendering conditional time-varying adjustments unreliable. ¹⁸ These models allow estimation of the 'average causal effect', which contrasts the counterfactual average outcome if everyone in the study population had been switched during the entire follow-up to the average outcome if everyone had not been switched during the entire follow-up. ¹³ In these models, we accounted for time dependent confounding of treatment exposure and censoring due to new relapses, cumulative relapse count, new T2 lesions on MRI, new gadolinium-enhancing lesions on MRI and updated EDSS status in 6-monthly intervals. We also accounted for time fixed baseline-confounding of EDSS status, disease duration, age and sex. The inverse probability weights were calculated as the product of the stabilised weights of treatment and censoring. For a more detailed methodology of the marginal structural model (MSM), see online supplementary appendix 1."	Methods	Methodological description of causal assumptions
	"The reported associations between switching therapy and treatment outcomes are a sum of a direct effect of the decision to switch therapy and the differences in effectiveness and pharmacokinetics of the therapies patients switched between. These two causal effects cannot be precisely quantified. However, in clinical practice, disentangling of these causal pathways is of less importance. The differences in outcomes seen between patients switching to TFL compared with DMF are likely due to this. This is corroborated by results from our previous nationwide cohort study comparing effectiveness in patients treated with either TFL or DMF. This suggested that DMF is more effective in preventing relapses and discontinuations due to activity, while the effects on disease worsening showed no difference. ¹¹ "	Discussion	Interpretation of results
Mesidor 2021 [24]	"To make a causal interpretation, assumptions such as conditional exchangeability, positivity, consistency, and accurate specification of the model should be respected. The associations of interest between interferon-beta/glatiramer acetate and the outcomes of MS were estimated on a dataset for which the distribution of potential confounders was balanced between the exposed and unexposed, reducing the likelihood of confounding. Although we controlled for many potential confounders such as number of relapses in the past 3 months, EDSS in the past 3 months, we did not have information on variables related to magnetic resonance imaging and neutralizing antibodies which can lead to residual confounding. Practical positivity can be violated due to small sample sizes in some intervals (Hernan and Robins, 2020). For this analysis, we considered people treated with interferon-beta/glatiramer acetate in the past 3 months versus those without medication in the past 3 months and, from our perspective, the exposure is well-defined. Finally, we assume that the weights were correctly specified (Cole and Hernan, 2008)."	Discussion	Interpretation of results
Palace 2019 [25]	"There are several other potential biases that must be considered before assuming a causal explanation for the DMTs that may have resulted in differences in the natural history of the two cohorts unrelated to treatment."	Discussion	Interpretation of results
Acyclic (n=1)	Zitat	Textstelle	Zusammenfassung
Mesidor 2021 [24]	"We hypothesized that the association between the use of interferon beta/glatiramer acetate and each of the three outcomes could be confounded by pregnancy in the past 3 months, number of relapses in the past 3 months, EDSS in the past 3 months, age and time since RRMS onset, as shown in the directed acyclic graphs (Fig. 1). Using the date of pregnancy, we created an indicator taking the value of 1 if the patient was pregnant in the past 3 months and 0 otherwise. The variables age and time since RRMS onset were created using the date of RRMS diagnosis and were both considered as continuous variables. We also considered the results of the cerebral spinal fluid (CSF) test and sex as predictors. A variable indicating the results of the CSF test was coded as follows: no test, normal test, abnormal test. An abnormal CSF test refers to positive oligoclonal bands."	Methods	Underlying assumption regarding the illustration of results

Abkürzungen: n = Number.

Teil II: Qualitative Interviews

Wir führten 8 Interviews mit Klinikern aus Deutschland (n=3) und der Schweiz (n=5) durch, die in Universitätskliniken (n=6) oder im ambulanten Bereich (n=2) arbeiteten, im Durchschnitt 20 Jahre (Interquartilsbereich [IQR] 17-23) und 19 Jahre (IQR 15-20) Berufserfahrung mit MS hatten und im letzten Jahr zwischen 0-550 Mal DMF und zwischen 0-30 Mal GA verschrieben hatten (ein Teilnehmer machte keine Angaben zur geschätzten Anzahl von Verordnungen). 6 von 8 Teilnehmern gaben an, dass bei der Verordnung von DMF- und GA-Leitlinien oder lokale Standards verwendet wurden.

Über die qualitative Interviewanalyse haben wir 35 Confounder identifiziert, die in der Literaturanalyse nicht identifiziert wurden. 18 potenzielle Confounder aus der Literaturanalyse wurden von den Klinikern bestätigt (Tabelle 2).

Teil III: Quantitativer Survey

Wir haben 70 Kliniker zur Teilnahme an der Umfrage eingeladen und 9 Personen aus unserem beruflichen Netzwerk gebeten, unsere Einladung weiterzuleiten (Anzahl der Einladungsempfänger unbekannt).

21 Personen begannen die Umfrage, von denen 7 sie beendeten. Diese 7 Teilnehmer machten soziodemografische Angaben und stammten aus Deutschland (n=1) und der Schweiz (n=6). Sie arbeiteten im Krankenhaus (n=5) und/oder im ambulanten Bereich (n=2) und/oder in einem anderen Bereich (n=1) und hatten eine mediane allgemeine Berufserfahrung von 14 Jahren (IQR 8-21) und eine mediane Berufserfahrung bei MS von 12 Jahren (IQR 9-14).

Von den 137 einzelnen potenziellen Confounder, die in der Umfrage genannt wurden, wurden die folgenden 40 als am wichtigsten für die Verordnung der Behandlungen eingestuft, mit einem Mittelwert von >4 auf einer Skala von "1 - keine Berücksichtigung" bis "5 - sehr starke Berücksichtigung" (Tabelle 4):

1. Pregnancy status²
2. Patient wish
3. Severity of relapses
4. Tolerance of subcutaneous injections
5. Willingness to inject a medication permanently

² Englisch war die Projektarbeitssprache. Auf die Übersetzung der Confounder und Confounder-Kategorien wurde verzichtet, um Übersetzungsfehler zu vermeiden.

6. Willingness to take pills permanently
7. Number of previous relapses (no defined index time)
8. Reason for discontinuing prior therapy
9. Time after natalizumab discontinuation
10. Absolute lymphocyte count; Lymphopenia
11. Drug efficacy (known or suspected)
12. Drug tolerance
13. MacDonald criteria; MS classification
14. Number of gadolinium enhancing lesions; Number of contrast-enhancing lesions;
Number of patients with gadolinium enhancing lesions; Patients with enhancement
15. Prognosis
16. Relapses in prior year
17. Side effects of other medications (multiple sclerosis and other indications)
18. Disease activity
19. Family planning
20. Pregnancy wish
21. Reason for dimethyl fumarate treatment
22. Relapses in the past 3 months
23. Relapses in the past 6 months
24. Remit of relapse
25. Timeliness of desire to have children
26. Disease severity at diagnosis
27. New lesions
28. Number of previous therapies
29. Patient concerns about drug side effects
30. Patient need for safety
31. Status of gadolinium enhancing T1 activity
32. Type of previous therapy
33. Clinical condition
34. Age at disease onset

35. Disease modifying therapy (DMT) reinitiation postpartum
36. Number of T2 lesions
37. Relapse in prior 3 years
38. Spinal lesions
39. Treatment naive; treatment history
40. Chronic infections

Von diesen 40 Confoundern wurden 50 % (n=20) nur in qualitativen Interviews, 35 % (n=14) nur in der Literatur und 15 % (n=6) in beiden identifiziert (Tabelle 4).

Im Rahmen des quantitativen Surveys wurden keine weiteren potenziellen Confounder identifiziert (Tabelle 2).

Tabelle 4: Wichtigkeit potenzieller Confounder

Bewertung der Wichtigkeit der individuellen potenziellen Confounder durch Kliniker, sortiert nach Mittelwert (hoch bis niedrig; 5-Punkte-Skala von '1 - keine Berücksichtigung' bis '5 - sehr starke Berücksichtigung').

Individuelle Confounder pro Kategorie (und alternative Ausdrücke, falls zutreffend)	Confounder-Kategorie	Source	n	Mean	SD	Median	Q1	Q2	Min	Max
Pregnancy status	Family planning, pregnancy, and birth	lit; int	7	4.86	0.35	5	5	5	4	5
Willingness to take pills permanently	Patient preferences, willingness, and concerns	int	7	4.71	0.70	5	5	5	3	5
Willingness to inject a medication permanently	Patient preferences, willingness, and concerns	int	7	4.71	0.45	5	4.5	5	4	5
Tolerance of subcutaneous injections	Tolerance and side effects	int	7	4.71	0.45	5	4.5	5	4	5
Severity of relapses	Relapse	int	7	4.71	0.45	5	4.5	5	4	5
Patient wish	Patient preferences, willingness, and concerns	int	7	4.71	0.45	5	4.5	5	4	5
Time after natalizumab discontinuation	Multiple sclerosis therapy	lit	6	4.67	0.47	5	4.25	5	4	5
Reason for discontinuing prior therapy	Multiple sclerosis therapy	lit	6	4.67	0.47	5	4.25	5	4	5
Number of previous relapses (no defined index time)	Relapse	lit; int	6	4.67	0.47	5	4.25	5	4	5
Side effects of other medications (multiple sclerosis and other indications)	Tolerance and side effects	int	7	4.57	0.49	5	4	5	4	5
Relapses in prior year	Relapse	lit	7	4.57	0.49	5	4	5	4	5
Prognosis	Prognosis	int	7	4.57	0.49	5	4	5	4	5
Number of gadolinium enhancing lesions; Number of contrast-enhancing lesions; Number of patients with gadolinium enhancing lesions; Patients with enhancement	Magnetic resonance imaging	lit; int	7	4.57	0.73	5	4.5	5	3	5
MacDonald criteria	Multiple sclerosis type and diagnosis criteria	lit	7	4.57	0.49	5	4	5	4	5
Drug tolerance	Tolerance and side effects	int	7	4.57	0.73	5	4.5	5	3	5
Drug efficacy (known or suspected)	Multiple sclerosis therapy	int	7	4.57	0.73	5	4.5	5	3	5
Absolute lymphocyte count or lymphopenia	Lab test	lit; int	7	4.57	0.49	5	4	5	4	5
Disease activity	Disease status (clinical and magnetic resonance imaging)	lit	8	4.50	0.50	4.5	4	5	4	5
Timeliness of desire to have children	Family planning, pregnancy, and birth	int	7	4.43	0.73	5	4	5	3	5
Remit of relapse	Relapse	int	7	4.43	0.73	5	4	5	3	5
Relapses in the past 6 months	Relapse	lit	7	4.43	0.73	5	4	5	3	5
Relapses in the past 3 months	Relapse	lit	7	4.43	0.73	5	4	5	3	5
Reason for dimethyl fumarate treatment	Multiple sclerosis therapy	lit	7	4.43	0.73	5	4	5	3	5
Pregnancy wish	Family planning, pregnancy, and birth	int	7	4.43	0.73	5	4	5	3	5
Family planning	Family planning, pregnancy, and birth	int	7	4.43	0.49	4	4	5	4	5
Disease severity at diagnosis	Disease status (clinical and magnetic resonance imaging)	int	8	4.38	0.48	4	4	5	4	5
Type of previous therapy	Multiple sclerosis therapy	lit; int	7	4.29	0.70	4	4	5	3	5
Status of gadolinium enhancing T1 activity	Magnetic resonance imaging	lit	7	4.29	0.70	4	4	5	3	5

Patient need for safety	Patient preferences, willingness, and concerns	int	7	4.29	1.39	5	4.5	5	1	5
Patient concerns about drug side effects	Patient preferences, willingness, and concerns	int	7	4.29	1.39	5	4.5	5	1	5
Number of previous therapies	Multiple sclerosis therapy	lit	7	4.29	0.70	4	4	5	3	5
New lesions	Magnetic resonance imaging	int	7	4.29	0.70	4	4	5	3	5
Clinical condition	Clinical status	int	8	4.25	0.43	4	4	4.25	4	5
Treatment naive or treatment history	Multiple sclerosis therapy	lit; int	7	4.14	0.99	4	4	5	2	5
Spinal lesions	Magnetic resonance imaging	int	7	4.14	0.64	4	4	4.5	3	5
Relapse in prior 3 years	Relapse	lit	7	4.14	1.36	5	4	5	1	5
Number of T2 lesions	Magnetic resonance imaging	lit	7	4.14	0.64	4	4	4.5	3	5
Disease modifying therapy (DMT) reinitiation postpartum	Family planning, pregnancy, and birth	lit	7	4.14	0.64	4	4	4.5	3	5
Age at disease onset	Socio-demographic characteristics	lit	7	4.14	0.64	4	4	4.5	3	5
Chronic infections	Comorbidity	int	8	4.13	0.60	4	4	4.25	3	5
Relapses in prior 2 years	Relapse	lit	7	4.00	1.31	4	4	5	1	5
Medical uncertainty with multiple sclerosis diagnosis	Multiple sclerosis type and diagnosis criteria	int	7	4.00	1.41	5	3.5	5	1	5
Gastrointestinal disorders	Tolerance and side effects	int	7	4.00	1.41	5	3.5	5	1	5
Duration of previous natalizumab use	Multiple sclerosis therapy	lit	7	4.00	0.76	4	3.5	4.5	3	5
Deterioration index (defined as the Expanded Disability Status Scale (EDSS) score divided by the disease duration)	Clinical status	lit	8	4.00	0.87	4	3	5	3	5
Comorbidities (general)	Comorbidity	lit; int	8	4.00	0.71	4	3.75	4.25	3	5
Expanded Disability Status Scale (EDSS) change preceding year	Clinical status	lit	8	3.88	1.27	4	4	5	1	5
Time on treatment	Multiple sclerosis therapy	lit	7	3.86	0.83	4	3	4.5	3	5
Time on first line treatment	Multiple sclerosis therapy	lit	7	3.86	0.83	4	3	4.5	3	5
Previous treatment with other disease modifying therapy (DMT)	Multiple sclerosis therapy	lit	7	3.86	0.83	4	3	4.5	3	5
Previous treatment with immunosuppressive treatment	Multiple sclerosis therapy	lit	7	3.86	0.99	4	3.5	4.5	2	5
Magnetic resonance imaging (MRI) brain lesions; number of lesions	Magnetic resonance imaging	lit; int	7	3.86	1.25	4	4	4.5	1	5
Duration of therapy	Multiple sclerosis therapy	lit	7	3.86	0.99	4	3.5	4.5	2	5
Duration of previous treatment therapy	Multiple sclerosis therapy	lit	7	3.86	0.99	4	3.5	4.5	2	5
Drug route of administration	Patient preferences, willingness, and concerns	int	7	3.86	1.36	4	3.5	5	1	5
Bowel symptoms	Patient-reported symptoms, health, functioning and well-being	lit	7	3.86	1.36	4	3.5	5	1	5
Age	Socio-demographic characteristics	lit; int	7	3.86	1.36	4	3.5	5	1	5
Steroid-treated relapses in prior year	Relapse	lit	6	3.83	1.34	4	4	4.75	1	5
Childhood multiple sclerosis diagnosis	Multiple sclerosis disease duration	lit	6	3.83	1.34	4	4	4.75	1	5
Time since last relapse	Relapse	lit	7	3.71	1.48	4	3	5	1	5
Previous treatment with glatiramer acetate	Multiple sclerosis therapy	lit	7	3.71	1.28	4	3.5	4.5	1	5
Lesion load	Magnetic resonance imaging	int	7	3.71	0.88	4	3.5	4	2	5
Age at onset of first multiple sclerosis symptoms	Multiple sclerosis disease duration	lit	7	3.71	1.39	4	3	5	1	5
Quality of life	Patient-reported symptoms, health, functioning and well-being	lit	7	3.57	1.50	4	2.5	5	1	5
Professional work activity	Lifestyle	int	7	3.57	0.90	4	3	4	2	5

Localisation of lesions	Magnetic resonance imaging	int	7	3.57	0.90	4	3	4	2	5
Breastfeeding	Family planning, pregnancy, and birth	lit; int	7	3.57	1.18	4	2.5	4.5	2	5
Neurological symptoms	Clinical status	int	8	3.50	1.00	4	3.75	4	1	4
Time since first multiple sclerosis symptoms	Multiple sclerosis disease duration	lit	7	3.43	1.18	4	3	4	1	5
Time since diagnosis of multiple sclerosis; disease duration	Multiple sclerosis disease duration	lit	7	3.43	1.18	4	3	4	1	5
Steroid-treated relapses in prior two years	Relapse	lit	7	3.43	1.40	4	2.5	4.5	1	5
Presence of at least 9 magnetic resonance imaging (MRI) T2 lesions and at least one contrast-enhancing MRI lesion	Magnetic resonance imaging	lit	7	3.43	1.18	4	3	4	1	5
Cognition status	Patient-reported symptoms, health, functioning and well-being	lit	7	3.43	1.05	4	3.5	4	1	4
Age at time of conception	Family planning, pregnancy, and birth	lit	7	3.43	1.18	4	3	4	1	5
Degree of pyramidal impairment	Clinical status	lit; int	8	3.38	1.32	3.5	2.75	4.25	1	5
Years of glatiramer acetate treatment	Multiple sclerosis therapy	lit	7	3.29	1.39	3	2.5	4.5	1	5
Time between symptom onset and disease modifying therapy (DMT) start	Multiple sclerosis disease duration	lit	7	3.29	1.28	4	2.5	4	1	5
Number of treated relapses	Relapse	lit; int	7	3.29	1.58	4	2	4.5	1	5
Human immunodeficiency virus (HIV)	Comorbidity	int	8	3.25	1.39	4	2.5	4	1	5
Disability status	Clinical status	lit; int	8	3.25	1.20	4	3	4	1	5
Volume of T2 lesions	Magnetic resonance imaging	lit	7	3.14	1.25	3	2.5	4	1	5
Size of lesions	Magnetic resonance imaging	int	7	3.14	1.25	3	2.5	4	1	5
Age of offspring at last follow-up	Family planning, pregnancy, and birth	lit	7	3.14	1.25	3	2.5	4	1	5
Year of treatment initiation	Multiple sclerosis therapy	lit	7	3.00	1.60	3	1.5	4.5	1	5
T1 hypointense lesion volume	Magnetic resonance imaging	lit	7	3.00	1.31	3	2	4	1	5
Magnetic resonance imaging (MRI) status	Magnetic resonance imaging	lit; int	7	3.00	1.51	4	1.5	4	1	5
Interferon beta treatment within past 3 months	Multiple sclerosis therapy	lit	7	3.00	1.60	3	1.5	4.5	1	5
Expanded Disability Status Scale (EDSS) in the past 3 months	Clinical status	lit	8	3.00	1.22	4	3	4	1	4
Enhancing lesion volume	Magnetic resonance imaging	lit	7	3.00	1.51	4	1.5	4	1	5
Depression (self-rated)	Patient-reported symptoms, health, functioning and well-being	lit	7	3.00	1.07	3	2.5	4	1	4
Expanded Disability Status Scale (EDSS) score (mean or median)	Clinical status	lit	8	2.88	1.17	3	2.5	4	1	4
Dermatological diseases	Comorbidity	int	8	2.88	0.93	3	2.75	3.25	1	4
Ambulation index	Clinical status	lit	8	2.88	1.62	3	1	4.25	1	5
Neurofilament light chain (NFL)	Lab test	int	7	2.86	1.46	3	1.5	4	1	5
Fatigue	Patient-reported symptoms, health, functioning and well-being	lit	7	2.86	0.99	3	2.5	3.5	1	4
Type and speciality of prescriber	Prescriber	lit	7	2.71	1.67	3	1	4	1	5
Previous treatment with interferon	Multiple sclerosis therapy	lit	7	2.71	1.48	2	1.5	4	1	5
Number of visits 1 year postpartum	Family planning, pregnancy, and birth	lit	7	2.71	1.28	2	2	3.5	1	5
Hepatitis	Comorbidity	int	8	2.63	1.32	3	1	4	1	4
Time since last visit	Multiple sclerosis therapy	lit	7	2.57	1.50	3	1	3.5	1	5
Numbness	Patient-reported symptoms, health, functioning and well-being	lit	7	2.57	1.40	2	1.5	3.5	1	5
Number of visits in pregnancy	Family planning, pregnancy, and birth	lit	7	2.57	1.18	2	2	3	1	5
Cerebrospinal fluid status	Lab test	lit; int	7	2.57	1.29	2	1.5	4	1	4

Cardiovascular comorbidities	Comorbidity	lit; int	8	2.50	1.41	2.5	1	3.25	1	5
Sexual Satisfaction Scale (SSS) score	Patient-reported symptoms, health, functioning and well-being	lit	7	2.43	1.50	2	1	3.5	1	5
Normalized white matter volume	Magnetic resonance imaging	lit	7	2.43	1.18	2	1.5	3.5	1	4
Immunoglobulin (IgM) in cerebrospinal fluid	Lab test	int	7	2.43	1.29	3	1	3.5	1	4
Gender	Socio-demographic characteristics	lit; int	7	2.43	0.90	2	2	3	1	4
Tuberculosis	Comorbidity	int	8	2.38	1.22	2.5	1	3.25	1	4
Patient Determined Disease Steps (PDDS) score	Patient-reported symptoms, health, functioning and well-being	lit	7	2.29	1.39	2	1	3	1	5
Overweight or obesity	Lifestyle	lit	7	2.29	1.03	2	1.5	3	1	4
Normalized grey matter volume	Magnetic resonance imaging	lit	7	2.29	1.16	2	1.5	3	1	4
Functional System (FS) score	Patient-reported symptoms, health, functioning and well-being	lit	7	2.29	1.16	2	1.5	3	1	4
Body mass index (BMI; height and weight)	Lifestyle	lit	7	2.29	1.03	2	1.5	3	1	4
Skin problems	Comorbidity	int	8	2.25	1.20	2	1	3.25	1	4
Brain volume; normalized brain volume	Magnetic resonance imaging	lit; int	6	2.17	0.90	2	2	2	1	4
Tobacco use	Lifestyle	lit	7	2.14	1.12	2	1	3	1	4
Multiple Sclerosis Impact Scale (MSIS) score	Patient-reported symptoms, health, functioning and well-being	lit	7	2.14	1.36	2	1	2.5	1	5
Multiple Sclerosis Impact Scale (MSIS) score	Patient-reported symptoms, health, functioning and well-being	lit	7	2.14	1.36	2	1	2.5	1	5
Modified Fatigue Impact Scale (MFIS) score	Patient-reported symptoms, health, functioning and well-being	lit	7	2.14	1.36	2	1	2.5	1	5
Weight (alone)	Lifestyle	lit	7	2.00	0.76	2	1.5	2.5	1	3
Hospitalisations in previous year	Comorbidity	lit	8	2.00	1.12	1.5	1	3	1	4
Depression (proxy assessment)	Clinical status	lit	8	2.00	1.00	2	1	2.25	1	4
Ethnicity or race	Socio-demographic characteristics	lit	7	1.86	0.83	2	1	2.5	1	3
Charlson Comorbidity Index (CCI) score	Comorbidity	lit	7	1.86	0.83	2	1	2.5	1	3
Previous outpatient care costs	Costs	lit	8	1.75	1.09	1	1	2.25	1	4
Multiple Sclerosis Functional Composite (MSFC) score	Clinical status	lit	8	1.75	1.09	1	1	2.25	1	4
Diabetes	Comorbidity	lit	8	1.75	1.09	1	1	2.25	1	4
Past steroid use	Multiple sclerosis therapy	lit	7	1.57	0.73	1	1	2	1	3
Past dalfampridine use	Multiple sclerosis therapy	lit	7	1.57	0.90	1	1	2	1	3
Educational level	Socio-demographic characteristics	lit	7	1.57	0.73	1	1	2	1	3
Previous hospital care costs	Costs	lit	8	1.50	1.00	1	1	1.25	1	4
Multiple Sclerosis Severity Score (MSSS)	Clinical status	lit	8	1.50	0.71	1	1	2	1	3
Within-country region	Geographical location and study center	lit	7	1.14	0.35	1	1	1	1	2
Region	Geographical location and study center	lit	7	1.14	0.35	1	1	1	1	2
Health plan type	Health insurance type and status	lit	7	1.14	0.35	1	1	1	1	2
Health insurance status	Health insurance type and status	lit	7	1.14	0.35	1	1	1	1	2
Country	Geographical location and study center	lit	7	1.14	0.35	1	1	1	1	2

Anmerkungen: Max = Maximum; Min = Minimum; n = Number of survey respondents; SD = Standard deviation; Q = Quartile.

Erläuterungen: Die Namen der individuellen potenziellen Confounder und Kategorien beziehen sich auf die im Survey Terminologie und können von der endgültigen Terminologie in Tabelle 2 abweichen.

Teil IV: Kategorisierung und Klassifizierung

Insgesamt identifizierten wir 160 potenzielle Confounder aus der Literatur (n=125; 78%) und qualitativen Interviews (n=35; 22%) und gruppieren sie in 20 Kategorien (Tabelle 2). Die 160 individuellen potenziellen Confounder wurden in 1-71 Quellen erwähnt (d.h. in Studienberichten und/oder qualitativen Interviews; Median 1; IQR 1-2). Die Kategorien, die die meisten individuellen potenziellen Confounder umfassten, standen im Zusammenhang mit der MS-Therapie (n=23), Patienten-berichteten Symptomen, Gesundheit, Funktionsfähigkeit und Wohlbefinden (n=21), der Magnetresonanztomographie (n=17), dem klinischen Status (n=15) und Komorbiditäten (n=14).

29 (18%) potenzielle Confounder wurden speziell in Bezug auf DMF und 54 (34%) in Bezug auf GA genannt; und 77 (48%) wurden sowohl in Bezug auf DMF als auch GA genannt.

12 (8%) potenzielle Confounder wurden speziell in Bezug auf das Outcome Behinderung genannt, 11 (7%) in Bezug auf die Schubrate und 4 (3%) in Bezug auf unerwünschte Ereignisse. 98 (61%) wurden in Bezug auf mehr als ein Outcome aus Behinderung, Schubrate, Lebensqualität und/oder unerwünschten Ereignissen genannt. Keiner der potenziellen Confounder wurde ausschließlich in Bezug auf die Lebensqualität genannt. 35 (22%) der potenziellen Confounder wurden ohne spezifischen Bezug zu einem der Outcomes aus Behinderung, Schubrate, Lebensqualität und/oder unerwünschten Ereignissen genannt (Tabelle 2).

148 (93%) potenzielle Confounder wurden als wahrscheinlich ein Prädiktor für die Intervention (DMF und GA) und 128 (80%) als wahrscheinlich auch ein Prädiktor für die Outcomes (Behinderung, Schubrate, Lebensqualität und unerwünschte Ereignisse) eingestuft; 11 (7%) potenzielle Confounder wurden als wahrscheinlich weder ein Prädiktor für die Intervention noch für die Outcomes (unklar) eingestuft; 1 (1%) potenzieller Confounder ("Type of MS") wurde nicht klassifiziert, da wir nur RRMS untersucht haben.

Die erwartete Richtung der Assoziation von potenziellen Confounder mit den Outcomes wurde meist als negativ (n=91; 57%) bewertet; d.h. je negativer die Bedeutung des Confounders für die Patienten, desto negativer die Auswirkungen auf die Outcomes aus ihrer Sicht [höhere Schubrate, schwerere Behinderung, geringere Lebensqualität, mehr unerwünschte Ereignisse] oder unklar (n=68; 43%), jedoch nie als positiv (n=0; 0%). Die Sicherheit der Bewertung der Richtung der Assoziation der 91 potenziellen Confounder mit den Outcomes basierte auf Expertenmeinungen; auch hier wurde 1 potenzieller Confounder ("Type of MS") nicht klassifiziert, da wir nur RRMS bewertet haben.

Die 160 individuellen potenziellen Confounder wurden schließlich zu 136 potenziellen Confounder geclustert (Tabelle 5).

Tabelle 5: Cluster potenzieller Confounder

Confounder, die auf ähnliche Weise gemessen oder bewertet werden können, alphabetisch und nach Kategorie der Confounder sortiert.

Confounder-Cluster pro Kategorie	Individuelle potenzielle Confounder pro Kategorie (und alternative Ausdrücke, falls zutreffend)
Clinical status	Clinical status
Ambulation index	Ambulation index
Clinical condition	Clinical condition
Depression diagnosis	Depression diagnosis
Deterioration index (defined as the Expanded Disability Status Scale (EDSS) score divided by the disease duration)	Deterioration index (defined as the Expanded Disability Status Scale (EDSS) score divided by the disease duration)
Disability	Disability status
Expanded Disability Status Scale (EDSS)	Expanded Disability Status Scale (EDSS) change preceding year
	Expanded Disability Status Scale (EDSS) in the past 3 months
	Expanded Disability Status Scale (EDSS) score (mean or median)
	Expanded Disability Status Scale (EDSS) scores/improvement at the end of the study treatment/last check-up
Hamilton Anxiety Rating Scale (HAM-A)	Hamilton Anxiety Rating Scale (HAM-A)
Hamilton Depression Rating Scale (HAM-D)	Hamilton Depression Rating Scale (HAM-D)
Multiple Sclerosis Functional Composite (MSFC) score	Multiple Sclerosis Functional Composite (MSFC) score
Multiple Sclerosis Severity Score (MSSS)	Multiple Sclerosis Severity Score (MSSS)
Neurological symptoms	Neurological symptoms
Pyramidal impairment	Pyramidal impairment
Comorbidity	Comorbidity
Allergy	Allergy
Cardiovascular comorbidities	Cardiovascular comorbidities
Charlson Comorbidity Index (CCI) score	Charlson Comorbidity Index (CCI) score
Chronic infections	Chronic infections
Comorbidities (general)	Comorbidities (general)
Dermatological diseases	Dermatological diseases
Diabetes	Diabetes
Headache	Headache
Hepatitis	Hepatitis
Hospitalisations in previous year	Hospitalisations in previous year
Human immunodeficiency virus (HIV)	Human immunodeficiency virus (HIV)
Skin problems	Skin problems
Thyroid dysfunction	Thyroid dysfunction
Tuberculosis	Tuberculosis
Costs	Costs
Previous hospital care costs	Previous hospital care costs
Previous outpatient care costs	Previous outpatient care costs
Disease status (clinical and magnetic resonance imaging)	Disease status (clinical and magnetic resonance imaging)
Disease activity	Disease activity
Disease course	Disease course

Disease severity	Disease severity
Disease severity at diagnosis	Disease severity at diagnosis
Family planning, pregnancy, and birth	Family planning, pregnancy, and birth
Age at time of conception	Age at time of conception
Age of offspring at last follow-up	Age of offspring at last follow-up
Breastfeeding	Breastfeeding
Disease modifying therapy (DMT) reinitiation postpartum	Disease modifying therapy (DMT) reinitiation postpartum
Family planning	Family planning
Number of visits 1 year postpartum	Number of visits 1 year postpartum
Number of visits in pregnancy	Number of visits in pregnancy
Pregnancy in the past 3 months	Pregnancy in the past 3 months
Pregnancy status	Pregnancy status
Pregnancy wish	Pregnancy wish
Timeliness of desire to have children	Timeliness of desire to have children
Geographical location and study center	Geographical location and study center
Country	Country
Region	Region
	Within-country region
Study center; Multiple sclerosis (MS) center	Study center; Multiple sclerosis (MS) center
Health insurance type and status	Health insurance type and status
Health insurance status	Health insurance status
Health plan type	Health plan type
Lab test	Lab test
Absolute lymphocyte count; Lymphopenia	Absolute lymphocyte count; Lymphopenia
Cerebrospinal fluid status	Cerebrospinal fluid status
Immunoglobulin (IgM) in cerebrospinal fluid	Immunoglobulin (IgM) in cerebrospinal fluid
Neurofilament light chain (NFL)	Neurofilament light chain (NFL)
Lifestyle	Lifestyle
Height and weight	Body mass index (BMI); or height and weight
	Overweight or obesity
	Weight
Tobacco use	Tobacco use
Type of professional work activity	Type of professional work activity
Magnetic resonance imaging	Magnetic resonance imaging
Brain volume; normalized brain volume	Brain volume; normalized brain volume
Enhancing lesion volume	Enhancing lesion volume
Lesion load	Lesion load
Localisation of lesions	Localisation of lesions
Magnetic resonance imaging (MRI) brain lesions; number of lesions	Magnetic resonance imaging (MRI) brain lesions; number of lesions
Magnetic resonance imaging (MRI) status	Magnetic resonance imaging (MRI) status
New lesions	New lesions
Normalized grey matter volume	Normalized grey matter volume
Normalized white matter volume	Normalized white matter volume

Number of gadolinium enhancing lesions; Number of contrast-enhancing lesions; Number of patients with gadolinium enhancing lesions; Patients with enhancement	Number of gadolinium enhancing lesions; Number of contrast-enhancing lesions; Number of patients with gadolinium enhancing lesions; Patients with enhancement
Number of T2 lesions	Number of T2 lesions
Presence of at least 9 magnetic resonance imaging (MRI) T2 lesions and at least one contrast-enhancing MRI lesion	Presence of at least 9 magnetic resonance imaging (MRI) T2 lesions and at least one contrast-enhancing MRI lesion
Size of lesions	Size of lesions
Spinal lesions	Spinal lesions
Status of gadolinium enhancing T1 activity	Status of gadolinium enhancing T1 activity
T1 hypointense lesion volume	T1 hypointense lesion volume
Volume of T2 lesions	Volume of T2 lesions
Multiple sclerosis disease duration	Multiple sclerosis disease duration
Age at onset of first multiple sclerosis symptoms	Age at onset of first multiple sclerosis symptoms
Age at start of disease modifying therapy (DMT); Age at start of treatment	Age at start of disease modifying therapy (DMT); Age at start of treatment
Childhood multiple sclerosis diagnosis	Childhood multiple sclerosis diagnosis
Time between symptom onset and disease modifying therapy (DMT) start	Time between symptom onset and disease modifying therapy (DMT) start
Time since diagnosis of multiple sclerosis; disease duration	Time since diagnosis of multiple sclerosis; disease duration
Time since first multiple sclerosis symptoms	Time since first multiple sclerosis symptoms
Multiple sclerosis therapy	Multiple sclerosis therapy
Drug efficacy (known or suspected)	Drug efficacy (known or suspected)
Duration of previous natalizumab use	Duration of previous natalizumab use
Duration of previous treatment therapy	Duration of previous treatment therapy
Duration of therapy	Duration of therapy
Number of previous therapies	Number of previous therapies
Past dalfampridine use	Past dalfampridine use
Past steroid use	Past steroid use
Previous interferon treatment	Interferon beta treatment within past 3 months Previous treatment with interferon
Previous treatment	Previous treatment with other disease modifying therapy (DMT) Treatment naive; treatment history Type of previous therapy
Previous treatment with glatiramer acetate	Previous treatment with glatiramer acetate
Previous treatment with immunosuppressive therapy	Previous treatment with immunosuppressive therapy
Reason for dimethyl fumarate treatment	Reason for dimethyl fumarate treatment
Reason for discontinuing prior therapy	Reason for discontinuing prior therapy
Time after natalizumab discontinuation	Time after natalizumab discontinuation
Time on first line treatment	Time on first line treatment
Time on study	Time on study
Time on treatment	Time on treatment
Time since last visit	Time since last visit
Year of treatment initiation	Year of treatment initiation
Years of glatiramer acetate treatment before study	Years of glatiramer acetate treatment before study
Multiple sclerosis type and diagnosis criteria	Multiple sclerosis type and diagnosis criteria
MacDonald criteria; MS classification	MacDonald criteria; MS classification
Medical uncertainty with multiple sclerosis diagnosis	Medical uncertainty with multiple sclerosis diagnosis

Type of multiple sclerosis (MS) onset (monofocal or multifocal)	Type of multiple sclerosis (MS) onset (monofocal or multifocal)
Patient preferences, willingness, and concerns	Patient preferences, willingness, and concerns
Drug route of administration	Drug route of administration
Patient concerns about drug side effects	Patient concerns about drug side effects
Patient need for safety	Patient need for safety
Patient wish	Patient wish
Willingness to inject a medication permanently	Willingness to inject a medication permanently
Willingness to take pills permanently	Willingness to take pills permanently
Patient-reported symptoms, health, functioning and well-being	Patient-reported symptoms, health, functioning and well-being
Bowel symptoms	Bowel symptoms
Cognition (self-rated)	Cognition (self-rated)
Depression (self-rated)	Depression (self-rated)
EuroQoL-5D (EQ-5D)	EuroQoL-5D (EQ-5D)
Fatigue	Fatigue
Functional System (FS) score	Functional System (FS) score
Modified Fatigue Impact Scale (MFIS) score	Modified Fatigue Impact Scale (MFIS) score
Multiple Sclerosis Impact Scale (MSIS) score	Multiple Sclerosis Impact Scale (MSIS) physical subscale score
	Multiple Sclerosis Impact Scale (MSIS) psychological subscale score
Multiple Sclerosis International Questionnaire of Quality of Life (MusiQoL)	Multiple Sclerosis International Questionnaire of Quality of Life (MusiQoL)
Numbness	Numbness
Patient Determined Disease Steps (PDDS) score	Patient Determined Disease Steps (PDDS) score
Sexual Satisfaction Scale (SSS) score	Sexual Satisfaction Scale (SSS) score
Short Form-36 (SF-36) score	Short Form-36 (SF-36) bodily pain subscale score
	Short Form-36 (SF-36) general health subscale score
	Short Form-36 (SF-36) mental health subscale score
	Short Form-36 (SF-36) physical functioning subscale score
	Short Form-36 (SF-36) role emotional subscale score
	Short Form-36 (SF-36) role physical subscale score
	Short Form-36 (SF-36) social functioning subscale score
	Short Form-36 (SF-36) vitality subscale score
Prescriber	Prescriber
Type and speciality of prescriber	Type and speciality of prescriber
Prognosis	Prognosis
Prognosis	Prognosis
Relapse	Relapse
Relapse history	Number of treated relapses
	Number of previous relapses (no defined index time)
	Relapse in prior 3 years
	Relapses in prior 2 years
	Relapses in prior year
	Relapses in the past 3 months
	Relapses in the past 6 months
Remit of relapse	Remit of relapse
Severity of relapses	Severity of relapses

Steroid-treated relapses	Steroid-treated relapses in prior two years
	Steroid-treated relapses in prior year
Time since last relapse	Time since last relapse
Socio-demographic characteristics	Socio-demographic characteristics
Age	Age
Age at disease onset	Age at disease onset
Educational level	Educational level
Ethnicity or race	Ethnicity or race
Gender/sex	Gender/sex
Tolerance and side effects	Tolerance and side effects
Drug tolerance	Drug tolerance
Gastrointestinal disorders	Gastrointestinal disorders
Side effects of other medications (multiple sclerosis and other indications)	Side effects of other medications (multiple sclerosis and other indications)
Tolerance of subcutaneous injections	Tolerance of subcutaneous injections

Machbarkeitsanalyse

Für die Identifizierung der 160 Confounder wurden 62 Personentage (à 8 Arbeitsstunden) als notwendig erachtet (Projektplanung, Projektmanagement, Supervision und Ergebnisberichterstattung nicht berücksichtigt). Der arbeitsintensivste Teil war Teil I, Literaturanalyse (125 Confounder identifiziert; geschätzte 33 Personentage benötigt; Tabelle 6).

Tabelle 6: Geschätzter Zeitaufwand für die Identifizierung potenzieller Confounder

Teil der Studie *	Geschätzte Zeit (PT **)	n potenzielle Confounder
I (Literaturanalyse)	33	125
II (Qualitative Interviews)	16 ***	35
III (Quantitativer Survey)	8	Keine zusätzlichen potenziellen Confounder identifiziert
IV (Kategorisierung und Klassifizierung)	5	NA

Abkürzungen: n = Anzahl; NA = Nicht zutreffend; PT = Personentage.

Anmerkungen: * Die beinhaltenden Aufgaben sind dem Methodenteil zu entnehmen; ** Ein Personentag (PT) entspricht 8 Arbeitsstunden;

*** Die Arbeitszeit für die Transkription der Interviews wurde geschätzt, da diese extern vergeben wurde.

5 Diskussion

Mithilfe unseres präspezifizierten, strukturierten und schrittweisen Ansatzes, der sich auf Literatur und klinische Expertise stützt, haben wir 160 potenzielle Confounder identifiziert, die theoretisch bei der Bewertung der Auswirkungen von DMF und GA auf die jährliche Schubrate, das Fortschreiten der Behinderung, die gesundheitsbezogene Lebensqualität und unerwünschte Ereignisse bei Erwachsenen mit RRMS vordefiniert, bewertet und berücksichtigt werden könnten.

Die meisten Confounder wurden in der Literatur identifiziert, jedoch ergaben die Gespräche mit klinischen Experten 35 weitere Confounder. Damit wurden 22% der 160 potenziellen Confounder nie in statistischen Analysen der veröffentlichten Studien berücksichtigt. Zu diesen 35 gehörten alle potenziellen Confounder, die mit Präferenzen, Bereitschaft und Bedenken der Patienten zusammenhängen (z.B. Patientenwunsch, Sicherheitsbedürfnis oder Bereitschaft zur dauerhaften Einnahme von Medikamenten). Es ist also ein bemerkenswertes Ergebnis, dass solche patientenzentrierten Faktoren in keiner der statistischen Analysen in der Literatur berücksichtigt wurden. Darüber hinaus konnte auch von den 40 als am wichtigsten eingestuften Confoundern die Hälfte nicht in der Literatur und nur durch die Interviews mit klinischen Experten identifiziert werden, da sie in den veröffentlichten Studien nicht berücksichtigt wurden. Bemerkenswerterweise bezog sich jeder vierte dieser wichtigsten Confounder auf patientenorientierte Aspekte aus den Kategorien "Family planning, pregnancy, and birth" oder "Patient preferences, willingness, and concerns", was wiederum die mangelnde Berücksichtigung patientenzentrierter Faktoren, welche Behandlungsentscheidungen bestimmen, in den gängigen statistischen Analysen unterstreicht.

Für die überwiegende Mehrheit der identifizierten Faktoren (80 %; 128 von 160) haben Expertenbewertungen die formale Definition eines Confounders bestätigt, d.h. dass der Faktor sowohl mit der Intervention (Exposition) als auch mit dem Ergebnis (Outcome) assoziiert ist (eine allfällige Rolle als Mediator, d.h. als ein Faktor, der auf dem kausalen Pfad liegt, wurde nicht berücksichtigt).

Einige der potenziellen Confounder wurden in etlichen Studien berücksichtigt und in den Interviews häufig erwähnt (bis zu 71 Berücksichtigungen bzw. Nennungen). Zu den am häufigsten berücksichtigten potenziellen Confounder gehörten Alter, Geschlecht, EDSS-Score (Expanded Disability Status Scale), Krankheitsdauer und Anzahl der Schübe im letzten Jahr. Von diesen fünf Faktoren wurden jedoch nur zwei, die jeweils mit einem Schub zusammenhängen, von den klinischen Experten als sehr wichtig für die Verordnung von DMF und GA angesehen. Diese Ergebnisse bestätigen, was sich auch aus den Ergebnissen von Pufulete et al. [7] ableiten lässt, nämlich dass es Unterschiede zwischen der klinischen

Perspektive und der tatsächlichen statistischen Berücksichtigung potenzieller Confounder in Beobachtungsstudien gibt.

Insgesamt ist die Zahl von 160 einzelnen potenziellen Confounder für eine statistische Analyse dieser klinischen Fragestellung hoch [26, 27]. Eine solche Detailliertheit und Differenziertheit der Confounder und eine derartige Inkonsistenz der Definitionen in den verschiedenen Studien haben wir nicht erwartet. Um den Mangel an Standardisierung etwas abzumildern, haben wir die Faktoren unter Berücksichtigung ähnlicher Aspekte und/oder Messmethoden geclustert, so dass wir die Liste auf 136 potenzielle Confounder reduzieren konnten. Diese eher geringe Reduzierung verdeutlicht, dass es unter diesen mehr als 100 Faktoren, die eine Behandlungsentscheidung bestimmen können und mit der Patientenprognose zusammenhängen, tatsächlich zahlreiche eindeutig unterschiedliche konzeptionelle Aspekte gibt, die berücksichtigt werden müssen, um unverzerrte Schätzungen der Auswirkungen von Behandlungsentscheidungen zu erhalten.

Limitationen

Diese Studie hat eine Reihe von Limitationen.

Erstens umfasste die Stichprobe für die qualitativen Interviews und die Umfrage Experten aus Deutschland und der Schweiz mit unterschiedlichem Hintergrund und Ordnungsverhalten, und es handelte sich nicht um eine repräsentative Stichprobe aller Verordner von DMF oder GA. Die meisten der Befragten gaben an, dass sie, obwohl sie in spezialisierten MS-Zentren ein breites Spektrum von Patienten betreuen, nicht bzw. nicht mehr regelmäßig DMF oder GA verschreiben. Jedoch waren alle Teilnehmer mit den beiden Medikamenten, deren Verschreibungen und den entsprechenden Behandlungsentscheidungen vertraut, dies im Rahmen ihres beruflichen klinischen Hintergrunds mit einer durchschnittlichen Berufserfahrung von fast 20 Jahren im Bereich MS.

Zweitens könnte die hohe Zahl der in der Literatur identifizierten potenziellen Confounder auf unseren sehr sensitiven Ansatz zurückzuführen sein. Die Datenextraktion galt als gesättigt, wenn bei der Extraktion von 15 aufeinanderfolgenden Studien der zufällig ausgewählten Stichprobe keine zusätzlichen potenziellen Confounder mehr identifiziert wurden. Dies war mehr als die von Pufulete et al. [7] verwendete Sättigungsgröße von 10 Studien. Das Screening aller Publikationen seit 2020 und einer Zufallsstichprobe von Publikationen vor 2020 hatte die Absicht, den neueren Entwicklungen bei MS angemessenes Gewicht zu verleihen und neuere Literatur und neue Quellen von "real-world"-Daten intensiver zu erfassen. Im Vergleich zu Herz-Kreislauf-Erkrankungen beispielsweise, bei denen die Krankheitsmechanismen seit Jahrzehnten relativ gut bekannt sind (das Thema der Arbeit von Pufulete et al. [7]), entwickelt sich das Gebiet der MS viel schneller, und die Mechanismen, die zu Confounding-by-Indication führen, werden sich wahrscheinlich im Laufe der Zeit ändern. Diese Prozesse trugen zwar dazu

bei, eine vollständigere Liste potenzieller Confounder zu erstellen, waren jedoch sehr ressourcenintensiv.

Drittens wurde die Umfrage von den meisten der eingeladenen Freiwilligen nicht vollständig ausgefüllt, da die Liste der Confounder sehr lang war. Mehr Teilnehmer hätten möglicherweise ein differenzierteres Bild von der klinischen Relevanz der potenziellen Confounder geliefert und die Rangfolge nach der Bedeutung der Confounder leicht verändert, aber wir würden keine andere Interpretation der wichtigsten Ergebnisse erwarten.

Viertens stützten sich die qualitativen Interviews auf acht Experten und damit zwei weniger als beabsichtigt, da es Schwierigkeiten bei der Rekrutierung gab und wir den Eindruck gewonnen hatten, dass die Beiträge dieser Experten bereits gesättigt waren, und unsere Ergebnisse die Faktoren, die von Klinikern bei der Verordnung von DMF oder GA berücksichtigt werden, ausreichend widerspiegeln.

Fünftens basiert die Machbarkeitsbewertung auf der Zusammensetzung der methodischen und klinischen Erfahrung der Kliniker und Forscher unseres Teams und spezialisierten Zentrums, was mehrere Schritte erleichtert haben könnte. Möglicherweise haben wir daher den Zeitaufwand unterschätzt, der in anderen Settings oder bei anderen Themen erforderlich wäre.

Sechstens ist es für eine unverzerrte Bewertung von Behandlungseffekten in statistischen Analysen von nicht randomisierten Studien nicht nur entscheidend, dass der Confounder in der Analyse überhaupt berücksichtigt wird, sondern auch wie. Mit anderen Worten: Neben der Auswahl relevanter Confounder ist die Art und Weise, wie diese Confounder in der Analyse berücksichtigt werden, von entscheidender Bedeutung und führt zu zahlreichen analytischen Entscheidungen (sogenannten "analytical degrees of freedom"), die eine erhebliche Variation-of-effects zur Folge haben [28]. Dazu gehört, wie Confounder definiert und operationalisiert werden und wie sie in den Analysen verwendet werden (z.B. Matching, Stratifizierung oder Adjustierung). Mitunter ist es von entscheidender Bedeutung ob z.B. eine lineare oder eine J-förmige Beziehung angenommen wird [29], da dies die klinische Interpretation der Ergebnisse substanziell verändern und schwerwiegende Auswirkungen haben kann [30]. Trotz der entscheidenden Bedeutung dieses zusätzlichen Aspekts für eine potenzielle Target Trial Emulation lag die Erfassung solcher Informationen jedoch außerhalb des Rahmens dieses Projekts, müsste idealerweise allerdings auch beachtet werden.

Insgesamt liefert unsere Studie trotz dieser Einschränkungen und aller zu beachtenden Vorbehalte wichtige Erkenntnisse für die Verwendung dieses systematischen Ansatzes für eine Target Trial Emulation und die Glaubwürdigkeit der Ergebnisse von Target Trial Emulationen:

Erstens fehlt bei sämtlichen der systematisch analysierten statistischen Analysen in der biomedizinischen Literatur zu diesem Thema eine Adjustierung für einige der wichtigsten Confounder, denn patientenzentrierte Aspekte werden als Quelle von Confounding-by-indication in keinem einzigen untersuchten Fall berücksichtigt. Damit besteht die berechtigte Sorge, dass alle diese Analysen substanziell verzerrte Schätzungen für Behandlungseffekte liefern und nicht zuverlässig abschätzen lassen, welche Therapien welchen Nutzen und welchen Schaden aufweisen.

Zweitens lässt sich mit einem rein datengesteuerten Ansatz nicht die Annahme erfüllen, dass *alle wichtigen bekannten Confounder in den statistischen Modellen berücksichtigt werden*. Daher bietet die veröffentlichte Literatur nur beschränkte Anhaltspunkte für die Auswahl relevanter Confounder, und die vorhandenen Analysen spiegeln nicht die Behandlungsauswahl in der Praxis wider.

Drittens ist der datengesteuerte Ansatz sehr zeitaufwändig. In Anbetracht der begrenzten Effizienz bei der Identifizierung der relevanten Confounder hat dieser Ansatz ein schlechtes Kosten-Nutzen-Verhältnis.

Viertens, da die Hälfte der wichtigsten Confounder wahrscheinlich nicht oder zumindest nicht hinreichend vollständig in einer verfügbaren "real-world"-Datenbank enthalten sind, können solche Datenquellen keine hinreichend zuverlässige "real-world evidence" liefern, zumindest wenn keine Randomisierung eingesetzt wird, sondern nur statistische Verfahren zur Confounder-Kontrolle. Solange zur Verfügung stehende Datenquellen solche Informationen beispielsweise zu Patientenpräferenzen nicht enthalten, ist es naheliegend, dass auch künftige nicht-randomisierte Studien mit solchen Datenquellen das gleiche sehr niedrige Maß an Glaubwürdigkeit aufweisen.

Letzlich bleibt als Alternative die Durchführung einer randomisierten Studie, die in Routinedateninfrastrukturen eingebettet ist, und all diese Einschränkungen in Bezug auf Confounding Bias überwinden kann [31]. Dieses Konzept wäre eine machbare Alternative [32] zu einem Target Trial Ansatz, der darauf abzielt, einige der Confounding Biases statistisch zu berücksichtigen, wohl wissend, dass selbst zahlreiche bekannte Confounder nicht berücksichtigt werden, ganz zu schweigen von den unbekanntem Confoundern.

6 Fazit

Durch einen systematischen, literaturbasierten Ansatz und die Einbeziehung von klinischer Expertise haben wir viele individuelle potenzielle Confounder identifiziert, die bei der Bewertung der Auswirkungen von DMF und GA auf die jährliche Rückfallrate, das Fortschreiten der Behinderung, die gesundheitsbezogene Lebensqualität und unerwünschte Ereignisse bei Erwachsenen mit RRMS berücksichtigt werden sollten. Zahlreiche als besonders wichtig eingestufte Faktoren konnten nur durch Expertenbefragungen ermittelt werden, auch da patientenorientierte Faktoren in gegenwärtigen publizierten Analysen nicht berücksichtigt werden. Die Analyse bestätigt eine Diskrepanz zwischen häufig berücksichtigten Confoundern in Studien und klinischen Überlegungen, was eine differenzierte Adjustierung in zukünftigen nicht-randomisierten Studien nahelegt. Die Studie unterstreicht die Notwendigkeit, neben der Literatur auch klinische Expertise und patientenorientierte Faktoren systematisch einzubeziehen, um eine umfassende Bewertung potenzieller Confounder in Beobachtungsstudien zu gewährleisten.

Die Studie unterstreicht auch, dass selbst mit erheblichen Anstrengungen und einem systematischen Vorgehen ein Ansatz zur Target Trial Emulation derzeit höchstwahrscheinlich keine hinreichend zuverlässige "real-world evidence" liefern wird, denn es ist derzeit kaum möglich, zahlreiche wichtige Ursachen von Confoundingbias zu vermeiden.

7 Literatur

1. Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. *Lancet*. 2002;359:57–61. doi:10.1016/S0140-6736(02)07283-5.
2. Matthews AA, Danaei G, Islam N, Kurth T. Target trial emulation: applying principles of randomised trials to observational studies. *BMJ*. 2022;378:e071108. doi:10.1136/bmj-2022-071108.
3. Aronson JK, Bankhead C, Nunan D. Catalogue of Bias: Confounding. 2018. <https://catalogofbias.org/biases/confounding/#:~:text=A%20distortion%20that%20modifies%20an,the%20exposure%20and%20the%20outcome>. Accessed 14 Nov 2023.
4. Reeves BC, Deeks JJ, Higgins JPT, Shea B, Tugwell P, Wells GA. Including non-randomized studies on intervention effects. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. *Cochrane Handbook for Systematic Reviews of Interventions* version 6.4; 2023.
5. Malec SA, Wei P, Bernstam EV, Boyce RD, Cohen T. Using computable knowledge mined from the literature to elucidate confounders for EHR-based pharmacovigilance. *Journal of Biomedical Informatics*. 2021;117:103719. doi:10.1016/j.jbi.2021.103719.
6. VanderWeele TJ. Principles of confounder selection. *Eur J Epidemiol*. 2019;34:211–9. doi:10.1007/s10654-019-00494-6.
7. Pufulete M, Mahadevan K, Johnson TW, Pithara C, Redwood S, Benedetto U, et al. Confounders and co-interventions identified in non-randomized studies of interventions. *J Clin Epidemiol*. 2022;148:115–23. doi:10.1016/j.jclinepi.2022.03.018.
8. Gloy V, McLennan S, Rinderknecht M, Ley B, Meier B, Driessen S, et al. Uncertainties about the need for ethics approval in Switzerland: a mixed-methods study. *Swiss Med Wkly*. 2020;150:w20318. doi:10.4414/smw.2020.20318.
9. Bundesamt für Gesundheit. Humanforschungsgesetz und Ethikkommissionen für die Forschung: Faktenblatt der Koordinationsstelle Forschung am Menschen (kofam). 2016. https://www.kofam.ch/filemanager/downloads/gutachten_und_berichte/BAG_Faktenblatt_Ethik_A4_D_low.pdf. Accessed 12 Jul 2023.
10. Viswanathan M, Berkman ND, Dryden DM, Hartling L. *Assessing Risk of Bias and Confounding in Observational Studies of Interventions or Exposures: Further Development of the RTI Item Bank*. Rockville (MD); 2013.
11. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Alemtuzumab, Cladribin, Dimethylfumarat, Fingolimod, Natalizumab, Ocrelizumab, Ofatumumab, Ozanimod, Ponesimod und Teriflunomid zur Behandlung Erwachsener mit hochaktiver schubförmig remittierender multipler Sklerose: Projekt: A20-60. Vorbericht. Version 1.0. 2023. <https://www.iqwig.de/projekte/a20-60.html>. Accessed 11 Sep 2023.
12. U.S. National Library of Medicine. *Medical Subject Headings 2023: MeSH Browser*. 2023. <https://meshb.nlm.nih.gov/search>. Accessed 12 Jul 2023.
13. PubMed PubReMiner. 2023. <http://hgserver2.amc.nl/cgi-bin/miner/miner2.cgi>. Accessed 12 Jul 2023.
14. Wong SS-L, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. *JMLA*. 2006;94:451–5.
15. Waffenschmidt S, Navarro-Ruan T, Hobson N, Hausner E, Sauerland S, Haynes RB. Development and validation of study filters for identifying controlled non-randomized studies in PubMed and Ovid MEDLINE. *Res Syn Meth*. 2020;11:617–26. doi:10.1002/jrsm.1425.

16. Furlan AD, Irvin E, Bombardier C. Limited search strategies were effective in finding relevant nonrandomized studies. *J Clin Epidemiol.* 2006;59:1303–11. doi:10.1016/j.jclinepi.2006.03.004.
17. Kung J. Polyglot Search Translator. *J Can Health Libr Assoc.* 2022;43:35–9. doi:10.29173/jchla29600.
18. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372:n71. doi:10.1136/bmj.n71.
19. Universität Basel. SMSC - Swiss Multiple Sclerosis Cohort. 2023. <https://dkf.unibas.ch/de/kompetenzen/register-und-kohorten/swiss-ms-cohort/>. Accessed 12 Jul 2023.
20. Universität Basel. MultiSCRIPT - Behandlungsoptimierung bei Multipler Sklerose. 2022. <https://dkf.unibas.ch/de/kompetenzen/iict/multiscript/>. Accessed 12 Jul 2023.
21. Verbi. MAXQDA. 2023. <https://www.maxqda.com/>. Accessed 14 Aug 2023.
22. Tivian. EFS Survey. 2024. <https://www.unipark.com/en/>. Accessed 14 May 2024.
23. Buron MD, Kalincik T, Sellebjerg F, Sorensen PS, Magyari M. Effect of lateral therapy switches to oral moderate-efficacy drugs in multiple sclerosis: a nationwide cohort study. *J Neurol Neurosurg Psychiatry.* 2021;92:556–62. doi:10.1136/jnnp-2020-324869.
24. Mesidor M, Sylvestre MP, Rousseau MC, Roger E, Duquette P. A non-parametric propensity score for estimating the effect of interferon-beta or glatiramer acetate on long-term outcomes of multiple sclerosis. *Mult Scler Relat Disord.* 2021;53:103050. doi:10.1016/j.msard.2021.103050.
25. Palace J, Duddy M, Lawton M, Bregenzer T, Zhu F, Boggild M, et al. Assessing the long-term effectiveness of interferon-beta and glatiramer acetate in multiple sclerosis: final 10-year results from the UK multiple sclerosis risk-sharing scheme. *J Neurol Neurosurg Psychiatry.* 2019;90:251–60. doi:10.1136/jnnp-2018-318360.
26. Hemkens LG, Ewald H, Naudet F, Ladanie A, Shaw JG, Sajeev G, Ioannidis JPA. Interpretation of epidemiologic studies very often lacked adequate consideration of confounding. *J Clin Epidemiol.* 2018;93:94–102. doi:10.1016/j.jclinepi.2017.09.013.
27. Munkholm K, Faurholt-Jepsen M, Ioannidis JPA, Hemkens LG. Consideration of confounding was suboptimal in the reporting of observational studies in psychiatry: a meta-epidemiological study. *J Clin Epidemiol.* 2020;119:75–84. doi:10.1016/j.jclinepi.2019.12.002.
28. Vinatier C, Hoffmann S, Patel C, DeVito NJ, Cristea IA, Tierney B, et al. What is the vibration of effects? *BMJ Evid Based Med* 2024. doi:10.1136/bmjebm-2023-112747.
29. Wallach JD, Serghiou S, Chu L, Egilman AC, Vasiliou V, Ross JS, Ioannidis JPA. Evaluation of confounding in epidemiologic studies assessing alcohol consumption on the risk of ischemic heart disease. *BMC Med Res Methodol.* 2020;20:64. doi:10.1186/s12874-020-0914-6.
30. Chu L, Ioannidis JPA, Egilman AC, Vasiliou V, Ross JS, Wallach JD. Vibration of effects in epidemiologic studies of alcohol consumption and breast cancer risk. *Int J Epidemiol.* 2020;49:608–18. doi:10.1093/ije/dyz271.
31. Mc Cord KA, Al-Shahi Salman R, Treweek S, Gardner H, Strech D, Whiteley W, et al. Routinely collected data for randomized trials: promises, barriers, and implications. *Trials.* 2018;19:29. doi:10.1186/s13063-017-2394-5.
32. Janiaud P, Zecca C, Salmen A, Benkert P, Schädelin S, Orleth A, et al. MultiSCRIPT-Cycle 1-a pragmatic trial embedded within the Swiss Multiple Sclerosis Cohort (SMSC) on neurofilament light chain monitoring to inform personalized treatment decisions in multiple sclerosis: a study protocol for a randomized clinical trial. *Trials.* 2024;25:607. doi:10.1186/s13063-024-08454-6.

Anhang A Suchstrategien

Ovid MEDLINE(R) ALL 1946 to November 2, 2023 ([Search History Link](#))

Notes:

- Randomized trials study filter by Wong 2006 [14]: strategy minimizing difference between sensitivity and specificity
- Observational study filter by Waffenschmidt 2020 [15]: search filter with best specificity for controlled non-randomized studies (Ovid MEDLINE, adapted from PubMed)

#	Query	Hits
1	exp "multiple sclerosis, relapsing-remitting"/	8280
2	((multiple sclerosis or MS) and (relaps* or remitting*)).ti,ab.	19209
3	or/1-2	21068
4	exp "Dimethyl fumarate"/ or exp "Glatiramer acetate"/	2512
5	(Dimethyl fumarat* or Dimethylfumarat* or Tecfidera or Copaxone or Glatiramer*).ti,ab.	3558
6	or/4-5	4040
7	and/3,6	1842
8	exp Randomized controlled Trial/	603161
9	controlled clinical trial.pt.	95429
10	(randomized or placebo or randomly).ab.	1090823
11	clinical trials as topic.sh.	201375
12	trial.ti.	295595
13	exp epidemiologic studies/ or exp clinical trial/ or comparative study/	5290507
14	((control and study) or program).mp.	2676245
15	or/8-14	7436483
16	(animals/ not humans/) or comment/ or editorial/ or exp review/ or meta analysis/ or consensus/ or exp guideline/	9758095
17	hi.fs. or case report.mp.	720205
18	or/16-17	10384562
19	15 not 18	6002050
20	and/7,19	684

Embase 1974 to 2023 November 2, 2023 ([Search History Link](#))

Notes:

- Randomized trials study filter by Wong 2006 [14]: strategy minimizing difference between sensitivity and specificity
- Observational study filter by Furlan 2006 [16]: fixed method B for Embase

#	Query	Hits
1	*multiple sclerosis, relapsing-remitting/ or ((multiple sclerosis or MS) and (relaps* or remitting)).ti,ab.	41103
2	*Dimethyl fumarate/ or *Glatiramer/ or (Dimethyl fumarat* or Dimethylfumarat* or Tecfidera or Copaxone or Glatiramer*).ti,ab.	8286
3	and/1-2	4335
4	(random* or double-blind*).tw.	2040390
5	placebo*.mp.	526698
6	or/4-5 [Wong 2006]	2267896
7	Clinical article/ or controlled study/ or major clinical study/ or prospective study/ or cohort.mp. or compared.mp. or groups.mp. or multivariate.mp. [Furlan 2006: Fixed method B for Embase]	18211828
8	3 and (6 or 7)	3548
9	8 not medline.cr.	3505
10	9 not (exp animal/ not exp human/)	3413
11	10 not (Conference Abstract or Conference Review or Editorial).pt.	1295
12	remove duplicates from 11	1276
13	comment/ or editorial/ or exp review/ or meta analysis/ or consensus/	4116952
14	12 not 13	1031

Anhang B Interviewleitfaden

EINSTIEGSFRAGE

Welche Aspekte berücksichtigen Sie bei der Verschreibung von DMF und GA?

[ggf. nachfragen: inwiefern spielen weitere Aspekte (bspw. soziale, gesellschaftliche, ökonomische Aspekte) eine Rolle?]

KLINISCHE FALLBEISPIELE

Szenario 1

Eine 28-jährige, ansonsten gesunde Frau stellt sich mit einer akuten Optikusneuritis rechts vor. In der Anamnese wurde eine etwa dreiwöchige selbstlimitierte Episode von Kribbelparästhesien in der linken Hand vor zwei Jahren festgestellt. Das cMRT und der Liquor stützen die Diagnose einer RRMS. Während Methylprednisolon intravenös verabreicht wird, überlegen Sie, welche krankheitsmodifizierende Behandlung Sie angesichts des Wunsches der Patientin, bald schwanger zu werden, empfehlen sollen.

Welches Medikament würden Sie verschreiben und warum?

Inwiefern ziehen Sie die Verschreibung von DMF in Betracht oder warum nicht?

Inwiefern ziehen Sie die Verschreibung von GA in Betracht oder warum nicht?

[alternative Formulierungen: Wäre es auch möglich, DMF/GA zu verschreiben? Wenn ja, warum, wenn nein, warum nicht?; Welche Überlegungen sind bei der Verschreibung von DMF oder GA aus Ihrer Sicht zu berücksichtigen?]

Szenario 2

Ein 42-jähriger Mann mit RRMS ist seit der Diagnose vor sechs Jahren stabil mit EDSS 2.0 auf Interferon beta-1a eingestellt. Das jährliche Routine-cMRT zeigt drei neue Läsionen, eine davon kontrastmittelaufnehmend. Sie erwägen, eine Behandlungseskalation zu empfehlen. Im Labor zeigte sich eine leichte Lymphopenie von 700/mm³.

Welches Medikament würden Sie verschreiben und warum?

Inwiefern ziehen Sie die Verschreibung von DMF in Betracht oder warum nicht?

Inwiefern ziehen Sie die Verschreibung von GA in Betracht oder warum nicht?

[alternative Formulierungen: Wäre es auch möglich, DMF/GA zu verschreiben? Wenn ja, warum, wenn nein, warum nicht?; Welche Überlegungen sind bei der Verschreibung von DMF oder GA aus Ihrer Sicht zu berücksichtigen?]

Szenario 3

Eine 46-jährige Frau stellte sich letzte Woche mit akuter Schwäche im linken Bein in der Notaufnahme vor. Es wurde die Diagnose RRMS gestellt und drei Tage lang mit i.v. Methylprednisolon behandelt. Nun hat sie einen ambulanten Termin in Ihrer Sprechstunde, um die krankheitsmodifizierende Behandlung zu besprechen. Nach der Lektüre des Beipackzettels von DMF ist die Patientin besorgt über Nebenwirkungen, die ihre bereits bestehenden Darmprobleme verschlimmern könnten.

Welches Medikament würden Sie verschreiben und warum?

Inwiefern ziehen Sie die Verschreibung von DMF in Betracht oder warum nicht?

Inwiefern ziehen Sie die Verschreibung von GA in Betracht oder warum nicht?

[alternative Formulierungen: Wäre es auch möglich, DMF/GA zu verschreiben? Wenn ja, warum, wenn nein, warum nicht?; Welche Überlegungen sind bei der Verschreibung von DMF oder GA aus Ihrer Sicht zu berücksichtigen?]

Szenario 4

Bei einem 25-jährigen Mann wurde kürzlich RRMS diagnostiziert, nachdem er sich mit einem Kleinhirnsyndrom, konfluierenden periventrikulären Läsionen und mehreren Hirnstamm- und Kleinhirnläsionen im MRT vorgestellt hatte, einige davon kontrastmittelaufnehmend. Zu den Komorbiditäten gehören Fettleibigkeit, starkes Rauchen und unbehandelte leichte Psoriasis.

Welches Medikament würden Sie verschreiben und warum?

Inwiefern ziehen Sie die Verschreibung von DMF in Betracht oder warum nicht?

Inwiefern ziehen Sie die Verschreibung von GA in Betracht oder warum nicht?

[alternative Formulierungen: Wäre es auch möglich, DMF/GA zu verschreiben? Wenn ja, warum, wenn nein, warum nicht?; Welche Überlegungen sind bei der Verschreibung von DMF oder GA aus Ihrer Sicht zu berücksichtigen?]

Szenario 5

Bei einer 32-jährigen Frau wurde kürzlich RRMS diagnostiziert. Die Patientin hat chronische Hautirritationen.

Welches Medikament würden Sie verschreiben und warum?

Inwiefern ziehen Sie die Verschreibung von DMF in Betracht oder warum nicht?

Inwiefern ziehen Sie die Verschreibung von GA in Betracht oder warum nicht?

[alternative Formulierungen: Wäre es auch möglich, DMF/GA zu verschreiben? Wenn ja, warum, wenn nein, warum nicht?; Welche Überlegungen sind bei der Verschreibung von DMF oder GA aus Ihrer Sicht zu berücksichtigen?]

SOZIODEMOGRAPHISCHE ANGABEN

Arbeitsfeld(er) (Setting):

Berufliche Rolle(n):

Berufserfahrung (Jahre):

Davon Berufserfahrung in MS (Jahre):

Anzahl von Verschreibungen im vergangenen Jahr (DMF/GA):

Vorhandensein von Verschreibungsstandards (ja/nein):

Gebrauch von Verschreibungsstandards (ja/nein):

Anhang C Fragebogen



Universität
Basel

Departement
Klinische Forschung



Research Center for
Clinical Neuroimmunology
and Neuroscience Basel

Welcome to our survey on "Systematic confounder identification in the indication of relapsing-remitting multiple sclerosis (RRMS)"!

Thank you for contributing to the improvement of observational studies and real-world evidence in multiple sclerosis (MS). Factors influencing treatment decisions or individual risk profiles often act as confounders. Addressing such confounders in observational studies is of utmost importance.

We are conducting a study on systematic confounder identification in the indication of relapsing-remitting multiple sclerosis (RRMS) therapies, specifically dimethyl fumarate and glatiramer acetate, focusing on relapse rate, disability, quality of life, and adverse events.

Through literature analysis and expert interviews, we have identified numerous potential confounders.

We now seek your expert opinion to rate the relevance of these factors. Please indicate how much each factor influences your decision when prescribing dimethyl fumarate or glatiramer acetate on a 5-point scale from '1 - No consideration' to '5 - Very strong consideration'.

This survey takes approximately 20 minutes to complete.

The potential confounders fall into the following categories (ordered alphabetically):

- Clinical status
- Comorbidity
- Costs
- Disease status (clinical and magnetic resonance imaging)
- Family planning, pregnancy, and birth
- Health insurance type and status
- Lab test
- Lifestyle
- Geographical location
- Magnetic resonance imaging
- Multiple sclerosis duration
- Multiple sclerosis therapy
- Multiple sclerosis type and diagnosis criteria
- Patient preferences, willingness, and concerns
- Patient-reported outcome
- Prescriber
- Prognosis
- Relapse
- Socio-demographic characteristics
- Tolerance and side effects

At the end of the survey, you can add any other aspects you consider when prescribing dimethyl fumarate or glatiramer acetate.

This study is commissioned by the Foundation for Quality and Efficiency in Health Care (IQWiG) (GA23-02) and conducted by the Research Center for Clinical Neuroimmunology and Neuroscience Basel (RC2NB) at the University of Basel and the University Hospital Basel. Your responses will be anonymized and not linked to you. No ethics approval is required for this project. By participating, you consent to our use of your anonymized answers for data analysis. We intend to publish the results in a peer-reviewed medical journal to enhance MS research.

Thank you for your participation!

I did not take part in an interview related to this project.

I already took part in an interview related to this project.

11%

CONTINUE

Clinical status

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Ambulation index	<input type="radio"/>				
Clinical condition	<input type="radio"/>				
Degree of pyramidal impairment	<input type="radio"/>				
Depression (proxy assessment)	<input type="radio"/>				
Deterioration index (defined as the Expanded Disability Status Scale (EDSS) score divided by the disease duration)	<input type="radio"/>				
Disability status	<input type="radio"/>				
Expanded Disability Status Scale (EDSS) change preceding year	<input type="radio"/>				
Expanded Disability Status Scale (EDSS) in the past 3 months	<input type="radio"/>				
Expanded Disability Status Scale (EDSS) score (mean or median)	<input type="radio"/>				
Multiple Sclerosis Functional Composite (MSFC) score	<input type="radio"/>				
Multiple Sclerosis Severity Score (MSSS)	<input type="radio"/>				
Neurological symptoms	<input type="radio"/>				

Comorbidity

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Cardiovascular comorbidities	<input type="radio"/>				
Charlson Comorbidity Index (CCI) score	<input type="radio"/>				
Chronic infections	<input type="radio"/>				
Comorbidities (general)	<input type="radio"/>				
Dermatological diseases	<input type="radio"/>				
Diabetes	<input type="radio"/>				
Hepatitis	<input type="radio"/>				
Hospitalisations in previous year	<input type="radio"/>				
Human immunodeficiency virus (HIV)	<input type="radio"/>				
Skin problems	<input type="radio"/>				
Tuberculosis	<input type="radio"/>				

Costs

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Previous hospital care costs	<input type="radio"/>				
Previous outpatient care costs	<input type="radio"/>				

Disease status (clinical and magnetic resonance imaging, MRI)

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Disease activity	<input type="radio"/>				
Disease severity at diagnosis	<input type="radio"/>				

Family planning, pregnancy, and birth

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Age at time of conception	<input type="radio"/>				
Age of offspring at last follow-up	<input type="radio"/>				
Breastfeeding	<input type="radio"/>				
Disease modifying therapy (DMT) reinitiation postpartum	<input type="radio"/>				
Family planning	<input type="radio"/>				
Number of visits 1 year postpartum	<input type="radio"/>				
Number of visits in pregnancy	<input type="radio"/>				
Pregnancy wish	<input type="radio"/>				
Pregnancy status	<input type="radio"/>				
Timeliness of desire to have children	<input type="radio"/>				

Geographical location

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Country	<input type="radio"/>				
Region	<input type="radio"/>				
Within-country region	<input type="radio"/>				

Health insurance type and status

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Health insurance status	<input type="radio"/>				
Health plan type	<input type="radio"/>				

Lab test

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Absolute lymphocyte count or lymphopenia	<input type="radio"/>				
Cerebrospinal fluid status	<input type="radio"/>				
Immunoglobulin (IgM) in cerebrospinal fluid	<input type="radio"/>				
Neurofilament light chain (NFL)	<input type="radio"/>				

Lifestyle

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Body mass index (BMI; height and weight)	<input type="radio"/>				
Overweight or obesity	<input type="radio"/>				
Professional work activity	<input type="radio"/>				
Tobacco use	<input type="radio"/>				
Weight (alone)	<input type="radio"/>				

Magnetic resonance imaging (MRI)

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Magnetic resonance imaging (MRI) brain lesions; number of lesions	<input type="radio"/>				
Volume of T2 lesions	<input type="radio"/>				
Number of T2 lesions	<input type="radio"/>				
Number of gadolinium enhancing lesions; Number of contrast-enhancing lesions; Number of patients with gadolinium enhancing lesions; Patients with enhancement	<input type="radio"/>				
T1 hypointense lesion volume	<input type="radio"/>				
Brain volume; normalized brain volume	<input type="radio"/>				
Status of gadolinium enhancing T1 activity	<input type="radio"/>				
Presence of at least 9 magnetic resonance imaging (MRI) T2 lesions and at least one contrast-enhancing MRI lesion	<input type="radio"/>				
Enhancing lesion volume	<input type="radio"/>				
Normalized grey matter volume	<input type="radio"/>				
Normalized white matter volume	<input type="radio"/>				
New lesions	<input type="radio"/>				
Size of lesions	<input type="radio"/>				
Lesion load	<input type="radio"/>				
Localisation of lesions	<input type="radio"/>				
Magnetic resonance imaging (MRI) status	<input type="radio"/>				
Spinal lesions	<input type="radio"/>				

Multiple sclerosis disease duration

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Age at onset of first multiple sclerosis symptoms	<input type="radio"/>				
Childhood multiple sclerosis diagnosis	<input type="radio"/>				
Time between symptom onset and disease modifying therapy (DMT) start	<input type="radio"/>				
Time since diagnosis of multiple sclerosis; disease duration	<input type="radio"/>				
Time since first multiple sclerosis symptoms	<input type="radio"/>				

Multiple sclerosis therapy

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Drug efficacy (known or suspected)	<input type="radio"/>				
Duration of previous natalizumab use	<input type="radio"/>				
Duration of previous treatment therapy	<input type="radio"/>				
Duration of therapy	<input type="radio"/>				
Interferon beta treatment within past 3 months	<input type="radio"/>				
Number of previous therapies	<input type="radio"/>				
Past dalfampridine use	<input type="radio"/>				
Past steroid use	<input type="radio"/>				
Previous treatment with glatiramer acetate	<input type="radio"/>				
Previous treatment with immunosuppressive treatment	<input type="radio"/>				
Previous treatment with interferon	<input type="radio"/>				
Previous treatment with other disease modifying therapy (DMT)	<input type="radio"/>				
Reason for dimethyl fumarate treatment	<input type="radio"/>				
Reason for discontinuing prior therapy	<input type="radio"/>				
Time after natalizumab discontinuation	<input type="radio"/>				
Time on first line treatment	<input type="radio"/>				
Time on treatment	<input type="radio"/>				
Time since last visit	<input type="radio"/>				
Treatment naive or treatment history	<input type="radio"/>				
Type of previous therapy	<input type="radio"/>				
Year of treatment initiation	<input type="radio"/>				
Years of glatiramer acetate treatment	<input type="radio"/>				

Multiple sclerosis type and diagnosis criteria

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
MacDonald criteria	<input type="radio"/>				
Medical uncertainty with multiple sclerosis diagnosis	<input type="radio"/>				
Type of multiple sclerosis	<input type="radio"/>				

Patient preferences, willingness, and concerns

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Drug route of administration	<input type="radio"/>				
Patient concerns about drug side effects	<input type="radio"/>				
Patient need for safety	<input type="radio"/>				
Patient wish	<input type="radio"/>				
Willingness to inject a medication permanently	<input type="radio"/>				
Willingness to take pills permanently	<input type="radio"/>				

Patient-reported outcome

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Bowel symptoms	<input type="radio"/>				
Cognition status	<input type="radio"/>				
Depression (self-rated)	<input type="radio"/>				
Fatigue	<input type="radio"/>				
Functional System (FS) score	<input type="radio"/>				
Modified Fatigue Impact Scale (MFIS) score	<input type="radio"/>				
Multiple Sclerosis Impact Scale (MSIS) score	<input type="radio"/>				
Numbness	<input type="radio"/>				
Patient Determined Disease Steps (PDSS) score	<input type="radio"/>				
Quality of life	<input type="radio"/>				
Sexual Satisfaction Scale (SSS) score	<input type="radio"/>				

Prescriber

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Type and speciality of prescriber	<input type="radio"/>				

Prognosis

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Prognosis	<input type="radio"/>				

Relapse

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Number of previous relapses (no defined index time)	<input type="radio"/>				
Number of treated relapses	<input type="radio"/>				
Relapse in prior 3 years	<input type="radio"/>				
Relapses in prior 2 years	<input type="radio"/>				
Relapses in prior year	<input type="radio"/>				
Relapses in the past 3 months	<input type="radio"/>				
Relapses in the past 6 months	<input type="radio"/>				
Remit of relapse	<input type="radio"/>				
Severity of relapses	<input type="radio"/>				
Steroid-treated relapses in prior two years	<input type="radio"/>				
Steroid-treated relapses in prior year	<input type="radio"/>				
Time since last relapse	<input type="radio"/>				

Socio-demographic characteristics

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Age	<input type="radio"/>				
Age at disease onset	<input type="radio"/>				
Ethnicity or race	<input type="radio"/>				
Educational level	<input type="radio"/>				
Gender	<input type="radio"/>				

Tolerance and side effects

To what extent do you consider the following aspects when prescribing dimethyl fumarate or glatiramer acetate?

	No consideration				Very strong consideration
Drug tolerance	<input type="radio"/>				
Gastrointestinal disorders	<input type="radio"/>				
Side effects of other medications (multiple sclerosis and other indications)	<input type="radio"/>				
Tolerance of subcutaneous injections	<input type="radio"/>				

Additional factors

Please feel free to add any other aspects you consider when prescribing dimethyl fumarate or glatiramer acetate.

Separate multiple factors by comma.

Work experience

Please enter your post-graduation work experience as a physician in the text box below (in years).

... of which professional experience as a physician in the field of multiple sclerosis (in years):

Work setting

In which setting are you currently working?

- Hospital care Outpatient care (self-employed/employed) Rehabilitation clinic Other setting

Professional role

What is your current professional role?

- Physician resident (or equivalent) Physician specialist in a managerial position or senior physician (or equivalent)
- Physician specialist in a non-managerial position (or equivalent) Chief physician (or equivalent) Other role

Geographical location

Where are you currently mainly working?

- Austria
- Germany
- Switzerland
- Other European country
- Non-European country

Thank you for taking part in our survey!

You can now close this window.

100%

Anhang D Einschlussfähige Studienberichte

Erstautor	Jahr	Titel	Digital object identifier
AlRuthia, Y.	2021	Real-World Comparative Cost-Effectiveness Analysis of Different Classes of Disease-Modifying Therapies for Relapsing-Remitting Multiple Sclerosis in Saudi Arabia	10.3390/ijerph182413261
Alsop, J.	2017	Two studies in one: A propensity-score-matched comparison of fingolimod versus interferons and glatiramer acetate using real-world data from the independent German studies, PANGAEA and PEARL	10.1371/journal.pone.0173353
Amezcuca, L.	2023	Efficacy of Dimethyl Fumarate in Young Adults with Relapsing-Remitting Multiple Sclerosis: Analysis of the DEFINE, CONFIRM, and ENDORSE Studies	10.1007/s40120-023-00475-8
Bergvall, N.	2014	Relapse rates in patients with multiple sclerosis switching from interferon to fingolimod or glatiramer acetate: a US claims database study	10.1371/journal.pone.0088472
Boffa, G.	2020	Fingolimod and Dimethyl-Fumarate-Derived Lymphopenia is not Associated with Short-Term Treatment Response and Risk of Infections in a Real-Life MS Population	10.1007/s40263-020-00714-8
Boggild, M.	2009	Multiple sclerosis risk sharing scheme: two year results of clinical cohort study with historical comparator	10.1136/bmj.b4677
Bosco-Levy, P.	2022	Comparative effectiveness of dimethyl fumarate in multiple sclerosis	10.1111/bcp.15071
Boster, A.	2017	Comparative Effectiveness Research of Disease-Modifying Therapies for the Management of Multiple Sclerosis: Analysis of a Large Health Insurance Claims Database	10.1007/s40120-017-0064-x
Bovis, F.	2021	Treatment Response Score to Glatiramer Acetate or Interferon Beta-1a	10.1212/wnl.0000000000010991
Braune, S.	2018	Comparative effectiveness of delayed-release dimethyl fumarate versus interferon, glatiramer acetate, teriflunomide, or fingolimod: results from the German NeuroTransData registry	10.1007/s00415-018-9083-5
Buron, M. D.	2021	Effect of lateral therapy switches to oral moderate-efficacy drugs in multiple sclerosis: a nationwide cohort study	10.1136/jnnp-2020-324869
Ciplea, A. I.	2023	Safety evaluations of offspring breastfed by mothers receiving glatiramer acetate for relapsing multiple sclerosis	10.1016/j.msard.2023.104771
Cohan, S. L.	2018	Clinical outcomes in patients with relapsing-remitting multiple sclerosis who switch from natalizumab to delayed-release dimethyl fumarate: A multicenter retrospective observational study (STRATEGY)	10.1016/j.msard.2018.02.028
Comi, G.	2011	Phase III dose-comparison study of glatiramer acetate for multiple sclerosis	10.1002/ana.22316
Conde, S.	2019	Dimethyl fumarate and teriflunomide for multiple sclerosis in a real-life setting: a French retrospective cohort study	10.1111/ene.13839
Cree, B. A. C.	2020	Efficacy and Safety of 2 Fingolimod Doses vs Glatiramer Acetate for the Treatment of Patients With Relapsing-Remitting Multiple Sclerosis: A Randomized Clinical Trial	10.1001/jamaneurol.2020.2950
Cree, B. A. C.	2018	Phase IV study of retention on fingolimod versus injectable multiple sclerosis therapies: a randomized clinical trial	10.1177/1756286418774338
D'Amico, E.	2020	Dimethyl fumarate vs Teriflunomide: an Italian time-to-event data analysis	10.1007/s00415-020-09959-1
Diem, L.	2020	Dimethyl fumarate vs fingolimod following different pretreatments: A retrospective study	10.1212/nxi.0000000000000660
Diouf, I.	2023	Effectiveness of multiple disease-modifying therapies in relapsing-remitting multiple sclerosis: causal inference to emulate a multiarm randomised trial	10.1136/jnnp-2023-331499
Fernandez, O.	2017	Efficacy and Safety of Delayed-release Dimethyl Fumarate for Relapsing-remitting Multiple Sclerosis in Prior Interferon Users: An Integrated Analysis of DEFINE and CONFIRM	10.1016/j.clinthera.2017.06.012
Ferraro, D.	2018	First-line disease-modifying drugs in relapsing-remitting multiple sclerosis: an Italian real-life multicenter study on persistence	10.1080/03007995.2018.1451311
Filippi, M.	2006	Effects of oral glatiramer acetate on clinical and MRI-monitored disease activity in patients with relapsing multiple sclerosis: a multicentre, double-blind, randomised, placebo-controlled study	10.1016/s1474-4422(06)70327-1
Ford, C. C.	2022	Early versus delayed treatment with glatiramer acetate: Analysis of up to 27 years of continuous follow-up in a US open-label extension study	10.1177/13524585221094239
Gold, R.	2017	Long-term effects of delayed-release dimethyl fumarate in multiple sclerosis: Interim analysis of ENDORSE, a randomized extension study	10.1177/1352458516649037
Gold, R.	2012	Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis	10.1056/NEJMoa1114287

Gold, R.	2022	Long-term safety and efficacy of dimethyl fumarate for up to 13 years in patients with relapsing-remitting multiple sclerosis: Final ENDORSE study results	10.1177/13524585211037909
Gold, R.	2020	Safety and efficacy of delayed-release dimethyl fumarate in patients with relapsing-remitting multiple sclerosis: 9 years' follow-up of DEFINE, CONFIRM, and ENDORSE	10.1177/1756286420915005
Granqvist, M.	2018	Comparative Effectiveness of Rituximab and Other Initial Treatment Choices for Multiple Sclerosis	10.1001/jamaneurol.2017.4011
Granqvist, M.	2020	Comparative effectiveness of dimethyl fumarate as the initial and secondary treatment for MS	10.1177/1352458519866600
Guger, M.	2019	Real-life use of oral disease-modifying treatments in Austria	10.1111/ane.13097
Guger, M.	2020	Oral therapies for treatment of relapsing-remitting multiple sclerosis in Austria: a 2-year comparison using an inverse probability weighting method	10.1007/s00415-020-09811-6
Haas, J.	2005	Twenty-four-month comparison of immunomodulatory treatments - a retrospective open label study in 308 RRMS patients treated with beta interferons or glatiramer acetate (Copaxone)	10.1111/j.1468-1331.2005.00936.x
Hellwig, K.	2023	Glatiramer acetate or IFN-beta bridging therapy in women with relapsing multiple sclerosis planning a pregnancy	10.2217/nmt-2022-0032
Hersh, C. M.	2022	Real-world effectiveness of dimethyl fumarate versus fingolimod in a cohort of patients with multiple sclerosis using standardized, quantitative outcome metrics	10.1177/20552173211069852
Hillert, J.	2022	A comparative study of teriflunomide and dimethyl fumarate within the Swedish MS Registry	10.1177/13524585211019649
Iaffaldano, P.	2015	Fingolimod versus interferon beta/glatiramer acetate after natalizumab suspension in multiple sclerosis	10.1093/brain/awv260
Izquierdo, G.	2015	Effectiveness of glatiramer acetate compared to other multiple sclerosis therapies	10.1002/brb3.337
Johnson, K. P.	1996	Management of relapsing/remitting multiple sclerosis with copolymer 1 (Copaxone)	10.1177/135245859600100606
Johnson, K. P.	1998	Extended use of glatiramer acetate (Copaxone) is well tolerated and maintains its clinical effect on multiple sclerosis relapse rate and degree of disability. Copolymer 1 Multiple Sclerosis Study Group	10.1212/wnl.50.3.701
Khan, O.	2017	Efficacy and safety of a three-times-weekly dosing regimen of glatiramer acetate in relapsing-remitting multiple sclerosis patients: 3-year results of the Glatiramer Acetate Low-Frequency Administration open-label extension study	10.1177/1352458516664033
Khan, O. A.	2001	A prospective, open-label treatment trial to compare the effect of IFNbeta-1a (Avonex), IFNbeta-1b (Betaseron), and glatiramer acetate (Copaxone) on the relapse rate in relapsing-remitting multiple sclerosis: Results after 18 months of therapy	10.1191/135245801701567041
Lorscheider, J.	2021	Comparative analysis of dimethyl fumarate and fingolimod in relapsing-remitting multiple sclerosis	10.1007/s00415-020-10226-6
Lublin, F. D.	2017	Long-term follow-up of a randomized study of combination interferon and glatiramer acetate in multiple sclerosis: Efficacy and safety results up to 7 years	10.1016/j.msard.2017.09.012
Mesidor, M.	2021	A non-parametric propensity score for estimating the effect of interferon-beta or glatiramer acetate on long-term outcomes of multiple sclerosis	10.1016/j.msard.2021.103050
Miller, D. H.	2015	Effects of delayed-release dimethyl fumarate on MRI measures in the phase 3 CONFIRM study	10.1212/wnl.0000000000001360
Mitsikostas, D.	2022	Quality of life in people with multiple sclerosis receiving glatiramer acetate or interferon in Greek clinical practice	10.2217/nmt-2022-0004
Mori, M.	2019	Efficacy and safety of dimethyl fumarate in treatment-naive Japanese patients with multiple sclerosis: Interim analysis of the randomized placebo-controlled study	10.1177/2055217319852727
Muller, J.	2023	Comparative analysis of dimethyl fumarate and teriflunomide in relapsing-remitting multiple sclerosis	10.1111/ene.16044
Nehzat, N.	2021	Comparable Efficacy and Safety of Teriflunomide versus Dimethyl Fumarate for the Treatment of Relapsing-Remitting Multiple Sclerosis	10.1155/2021/6679197
O'Connor, P.	2009	250 microg or 500 microg interferon beta-1b versus 20 mg glatiramer acetate in relapsing-remitting multiple sclerosis: a prospective, randomised, multicentre study	10.1016/s1474-4422(09)70226-1
Palace, J.	2019	Assessing the long-term effectiveness of interferon-beta and glatiramer acetate in multiple sclerosis: final 10-year results from the UK multiple sclerosis risk-sharing scheme	10.1136/jnnp-2018-318360
Pantazou, V.	2021	First-ever treatment in multiple sclerosis	10.1016/j.neurol.2020.05.014
Pfeuffer, S.	2021	Teriflunomide treatment is associated with optic nerve recovery in early multiple sclerosis	10.1177/1756286421997372
Pilo de la Fuente, B.	2020	Three-Year Effectiveness of Dimethyl Fumarate in Multiple Sclerosis: A Prospective Multicenter Real-World Study	10.1007/s40263-020-00775-9

Reder, A. T.	2021	Real-world propensity score comparison of treatment effectiveness of peginterferon beta-1a vs. subcutaneous interferon beta-1a, glatiramer acetate, and teriflunomide in patients with relapsing-remitting multiple sclerosis	10.1016/j.msard.2021.102935
Rieckmann, P.	2021	Long-term efficacy and safety of three times weekly dosing regimen of glatiramer acetate in relapsing multiple sclerosis patients: Seven-year results of the Glatiramer Acetate Low-frequency Administration (GALA) open-label extension study	10.1177/20552173211061550
Rojas, J. I.	2022	Effectiveness and Safety of Early High-Efficacy Versus Escalation Therapy in Relapsing-Remitting Multiple Sclerosis in Argentina	10.1097/wnf.0000000000000503
Romeo, M. A. L.	2018	Assessing the role of innovative therapeutic paradigm on multiple sclerosis treatment response	10.1111/ane.12999
Rossi, S.	2014	Treatment options to reduce disease activity after natalizumab: paradoxical effects of corticosteroids	10.1111/cns.12282
Rovaris, M.	2003	Short-term correlations between clinical and MR imaging findings in relapsing-remitting multiple sclerosis	
Rovaris, M.	2001	Short-term brain volume change in relapsing-remitting multiple sclerosis: effect of glatiramer acetate and implications	10.1093/brain/124.9.1803
Saida, T.	2019	A randomized placebo-controlled trial of delayed-release dimethyl fumarate in patients with relapsing-remitting multiple sclerosis from East Asia and other countries	10.1186/s12883-018-1220-3
Salter, A.	2021	A propensity-matched comparison of long-term disability worsening in patients with multiple sclerosis treated with dimethyl fumarate or fingolimod	10.1177/17562864211021177
Sattarnezhad, N.	2022	Comparison of dimethyl fumarate and interferon outcomes in an MS cohort	10.1186/s12883-022-02761-8
Scott, T. F.	2021	Matching comparisons of therapeutic efficacy suggest better clinical outcomes for patients treated with peginterferon beta-1a than with glatiramer acetate	10.1177/1756286420975916
Signori, A.	2020	Cladribine vs other drugs in MS: Merging randomized trial with real-life data	10.1212/nxi.0000000000000878
Spelman, T.	2023	Comparative effectiveness of cladribine tablets versus other oral disease-modifying treatments for multiple sclerosis: Results from MSBase registry	10.1177/13524585221137502
Spelman, T.	2016	Comparative efficacy of first-line natalizumab vs IFN-beta or glatiramer acetate in relapsing MS	10.1212/cpj.0000000000000227
Stuchiner, T.	2020	Quality of life among injectable and oral disease-modifying therapy users in the Pacific Northwest Multiple Sclerosis Registry	10.1186/s12883-020-02016-4
Svenningsson, A.	2022	Safety and efficacy of rituximab versus dimethyl fumarate in patients with relapsing-remitting multiple sclerosis or clinically isolated syndrome in Sweden: a rater-blinded, phase 3, randomised controlled trial	10.1016/s1474-4422(22)00209-5
Viglietta, V.	2015	Efficacy of delayed-release dimethyl fumarate in relapsing-remitting multiple sclerosis: integrated analysis of the phase 3 trials	10.1002/acn3.148
Vollmer, T.	2008	Glatiramer acetate after induction therapy with mitoxantrone in relapsing multiple sclerosis	10.1177/1352458507085759
von Wyl, V.	2021	Disability progression in relapse-free multiple sclerosis patients on fingolimod versus interferon-beta/glatiramer acetate	10.1177/1352458520918489
Voskuhl, R. R.	2016	Estriol combined with glatiramer acetate for women with relapsing-remitting multiple sclerosis: a randomised, placebo-controlled, phase 2 trial	10.1016/s1474-4422(15)00322-1
Wolinsky, J. S.	2001	United States open-label glatiramer acetate extension trial for relapsing multiple sclerosis: MRI and clinical correlates. Multiple Sclerosis Study Group and the MRI Analysis Center	10.1177/135245850100700107
Wolinsky, J. S.	2001	United States open-label glatiramer acetate extension trial for relapsing multiple sclerosis: MRI and clinical correlates	10.1191/135245801667520627
Wolinsky, J. S.	2015	GLACIER: An open-label, randomized, multicenter study to assess the safety and tolerability of glatiramer acetate 40 mg three-times weekly versus 20 mg daily in patients with relapsing-remitting multiple sclerosis	10.1016/j.msard.2015.06.005
Yeh, W. Z.	2021	Natalizumab, Fingolimod, and Dimethyl Fumarate Use and Pregnancy-Related Relapse and Disability in Women With Multiple Sclerosis	10.1212/wnl.0000000000012084
Zhu, C.	2023	Comparison Between Dimethyl Fumarate, Fingolimod, and Ocrelizumab After Natalizumab Cessation	10.1001/jamaneurol.2023.1542

Anhang E Ausschlussgründe für Studienberichte

Erstautor	Jahr	Titel	Digital object identifier	Ausschlussgrund
AbdelRazek, M. A.	2023	Exploring the effect of glatiramer acetate on cerebral gray matter atrophy in multiple sclerosis	10.1016/j.jns.2022.120501	Outcome
Agrawal, A.	2023	A Real-World Experience of Azathioprine Versus First-Line Disease-Modifying Therapy in Relapsing-Remitting Multiple Sclerosis-A Prospective Cohort Study	10.3390/brainsci13091249	No confounder adjustment
Al Turaiki, A. M.	2019	Assessment of safety and effectiveness of oral multiple sclerosis medication	10.15537/smj.2019.11.24630	No confounder adjustment
Al-Abri, S.	2023	Initial clinical and radiological features of patients with multiple sclerosis in Oman	10.1016/j.msard.2023.104734	Population
Al-iedani, O.	2022	Neurometabolic changes in multiple sclerosis: Fingolimod versus beta interferon or glatiramer acetate therapy	10.1111/jon.13032	Population
Allan, M.	2020	A Retrospective Analysis of Real-World Discontinuation Rates with Delayed-Release Dimethyl Fumarate in Patients with Relapsing-Remitting Multiple Sclerosis	10.1007/s40120-019-00174-3	No confounder adjustment
Alping, P.	2021	Safety of Alemtuzumab and Autologous Hematopoietic Stem Cell Transplantation Compared to Noninduction Therapies for Multiple Sclerosis	10.1212/wnl.0000000000011545	Outcome
Bawand, R.	2022	Effects of disease-modifying treatments discontinuation in patients with relapsing-remitting multiple sclerosis: A 5 year prospective cohort study	10.1016/j.msard.2022.103857	No confounder adjustment
Bergamaschi, R.	2016	Immunomodulatory therapies delay disease progression in multiple sclerosis	10.1177/1352458512445941	Outcome
Boiko, A. N.	2013	Efficacy and tolerance of glatiramer acid (Copaxone) on long-term use: 10 years of experience at the moscow city multiple sclerosis center	10.1007/s11055-013-9773-5	No confounder adjustment
Caon, C.	2006	Clinical course after change of immunomodulating therapy in relapsing-remitting multiple sclerosis	10.1111/j.1468-1331.2006.01273.x	No confounder adjustment
Carra, A.	2003	A retrospective, observational study comparing the four available immunomodulatory treatments for relapsing-remitting multiple sclerosis	10.1046/j.1468-1331.2003.00669.x	No confounder adjustment
Chan, A.	2017	Comparative effectiveness of delayed-release dimethyl fumarate versus glatiramer acetate in multiple sclerosis patients: results of a matching-adjusted indirect comparison	10.2217/cer-2016-0085	Study design
Cheshmavar, M.	2021	Rituximab and glatiramer acetate in secondary progressive multiple sclerosis: A randomized clinical trial	10.1111/ane.13344	Population
China, A.	2020	Real-World Safety and Effectiveness of Dimethyl Fumarate in Hispanic or Latino Patients with Multiple Sclerosis: 3-Year Results from ESTEEM	10.1007/s40120-020-00192-6	Study design
Cinar, B. P.	2017	Cognitive dysfunction in patients with multiple sclerosis treated with first-line disease-modifying therapy: a multi-center, controlled study using the BICAMS battery	10.1007/s10072-016-2775-7	Outcome
Ciplea, A. I.	2022	Eighteen-month safety analysis of offspring breastfed by mothers receiving glatiramer acetate therapy for relapsing multiple sclerosis - COBRA study	10.1177/13524585221083982	Outcome
Clerico, M.	2014	Treatment of relapsing-remitting multiple sclerosis after 24 doses of natalizumab: evidence from an Italian spontaneous, prospective, and observational study (the TY-STOP Study)	10.1001/jamaneurol.2014.1200	Intervention
Coles, A. J.	2012	Alemtuzumab for patients with relapsing multiple sclerosis after disease-modifying therapy: a randomised controlled phase 3 trial	10.1016/s0140-6736(12)61768-1	Intervention
Comi, G.	2009	Effect of glatiramer acetate on conversion to clinically definite multiple sclerosis in patients with clinically isolated syndrome (PreCISe study): a randomised, double-blind, placebo-controlled trial	10.1016/s0140-6736(09)61259-9	Population
Comi, G.	2001	European/Canadian multicenter, double-blind, randomized, placebo-controlled study of the effects of glatiramer acetate on magnetic resonance imaging-measured disease activity and burden in patients with relapsing multiple sclerosis	10.1002/ana.64	Outcome
Comi, G.	2001	European/Canadian multicenter, double-blind, randomized, placebo-controlled study of the effects of glatiramer acetate on magnetic resonance imaging--measured disease activity and burden in patients with relapsing multiple sclerosis. European/Canadian Glatiramer Acetate Study Group		Outcome
Comi, G.	2023	Impact of treatment with dimethyl fumarate on sleep quality in patients with relapsing-remitting multiple sclerosis: A multicentre Italian wearable tracker study	10.1177/20552173221144229	No confounder adjustment

Cutter, G.	2019	Higher satisfaction and adherence with glatiramer acetate 40mg/mL TIW vs 20mg/mL QD in RRMS	10.1016/j.msard.2019.04.036	Outcome
Cutter, G.	2020	Satisfaction and adherence with glatiramer acetate 40mg/mL TIW in RRMS after 12 months, and the effect of switching from 20mg/mL QD	10.1016/j.msard.2020.101957	Outcome
D'Amico, E.	2021	Injectable Versus Oral First-Line Disease-Modifying Therapies: Results from the Italian MS Register	10.1007/s13311-020-01001-6	Outcome
Debouverie, M.	2007	A longitudinal observational study of a cohort of patients with relapsing-remitting multiple sclerosis treated with glatiramer acetate	10.1111/j.1468-1331.2007.01964.x	No confounder adjustment
Dobson, R.	2022	Evaluating the feasibility of a real world pharmacovigilance study (OPTIMISE:MS)	10.1016/j.msard.2022.103894	No confounder adjustment
Earla, J. R.	2021	Comparative treatment effectiveness of oral fingolimod and conventional injectable disease-modifying agents in multiple sclerosis	10.1002/phar.2517	Outcome
Freedman, M. S.	2015	A randomized trial of teriflunomide added to glatiramer acetate in relapsing multiple sclerosis	10.1177/2055217315618687	Intervention
Gasperini, C.	2021	Scoring the 10-year risk of ambulatory disability in multiple sclerosis: the RoAD score	10.1111/ene.14845	Study design
Glasmacher, S. A.	2022	The influence of disease-modifying therapy on hidden disability burden in people with newly diagnosed relapsing-remitting multiple sclerosis	10.1016/j.msard.2022.103837	Outcome
Greenberg, B.	2020	Multiple sclerosis relapse rates and healthcare costs of two versions of glatiramer acetate	10.1080/03007995.2020.1760808	Population
Guger, M.	2023	Effects of horizontal versus vertical switching of disease-modifying treatment after platform drugs on disease activity in patients with relapsing-remitting multiple sclerosis in Austria	10.1007/s00415-023-11644-y	Outcome
Hanninen, K.	2022	Initial treatment strategy and clinical outcomes in Finnish MS patients: a propensity-matched study	10.1007/s00415-021-10673-9	Outcome
Harding, K.	2019	Clinical Outcomes of Escalation vs Early Intensive Disease-Modifying Therapy in Patients With Multiple Sclerosis	10.1001/jamaneurol.2018.4905	Intervention
Healy, B. C.	2010	A method for evaluating treatment switching criteria in multiple sclerosis	10.1177/1352458510379245	Study design
Hersh, C. M.	2017	Comparative efficacy and discontinuation of dimethyl fumarate and fingolimod in clinical practice at 24-month follow-up	10.1177/205521731715485	Population
Hersh, C. M.	2020	Effect of switching from natalizumab to moderate-vs high-efficacy DMT in clinical practice	10.1212/cpj.0000000000000809	Population
Hewes, D.	2017	SIRT1 as a potential biomarker of response to treatment with glatiramer acetate in multiple sclerosis	10.1016/j.yexmp.2017.01.014	Outcome
Hojsgaard Chow, H.	2023	Dimethyl fumarate treatment of primary progressive multiple sclerosis: results of an open-label extension study	10.1016/j.msard.2022.104458	Population
Honce, J. M.	2020	Brain Atrophy Rates for Stable Multiple Sclerosis Patients on Long-Term Fingolimod versus Glatiramer Acetate	10.3389/fneur.2020.01045	No confounder adjustment
Hou, J.	2021	Comparison of Dimethyl Fumarate vs Fingolimod and Rituximab vs Natalizumab for Treatment of Multiple Sclerosis	10.1001/jamanetworkopen.2021.34627	Population
Iaffaldano, P.	2021	Long-term disability trajectories in relapsing multiple sclerosis patients treated with early intensive or escalation treatment strategies	10.1177/17562864211019574	Intervention
Izquierdo, G.	2017	The real-world effectiveness and safety of fingolimod in relapsing-remitting multiple sclerosis patients: An observational study	10.1371/journal.pone.0176174	Intervention
Johnson, B. H.	2015	Platform Therapy Compared with Natalizumab for Multiple Sclerosis: Relapse Rates and Time to Relapse Among Propensity Score-Matched US Patients	10.1007/s40263-015-0251-1	Population
Jongen, P. J.	2017	Glatiramer acetate treatment persistence - but not adherence - in multiple sclerosis patients is predicted by health-related quality of life and self-efficacy: a prospective web-based patient-centred study (CAIR study)	10.1186/s12955-017-0622-z	Study design
Jongen, P. J.	2014	Fatigue and health-related quality of life in relapsing-remitting multiple sclerosis after 2 years glatiramer acetate treatment are predicted by changes at 6 months: an observational multi-center study	10.1007/s00415-014-7363-2	No confounder adjustment
Jongen, P. J.	2010	Health-related quality of life in relapsing remitting multiple sclerosis patients during treatment with glatiramer acetate: a prospective, observational, international, multi-centre study	10.1186/1477-7525-8-133	No confounder adjustment
Jozef, M.	2023	Medication adherence and health outcomes in persons with multiple sclerosis treated with dimethyl fumarate	10.1016/j.msard.2023.104615	Study design
Kalincik, T.	2021	Effect of Disease-Modifying Therapy on Disability in Relapsing-Remitting Multiple Sclerosis Over 15 Years	10.1212/wnl.0000000000011242	Population
Kalincik, T.	2015	Switch to natalizumab versus fingolimod in active relapsing-remitting multiple sclerosis	10.1002/ana.24339	Intervention

Karampampa, K.	2023	Disease-modifying therapies and cost-of-illness progression among people newly diagnosed with multiple sclerosis: a national register-based cohort study covering treatment initiation with interferons, glatiramer acetate or natalizumab	10.1136/bmjopen-2022-067516	Population
Koch, M. W.	2022	Early first-line treatment response and subsequent disability worsening in relapsing-remitting multiple sclerosis	10.1111/ene.15220	Study design
Koeditz, D.	2022	Comparing the long-term clinical and economic impact of ofatumumab versus dimethyl fumarate and glatiramer acetate in patients with relapsing multiple sclerosis: A cost-consequence analysis from a societal perspective in Germany	10.1177/20552173221085741	Study design
Kremenchutzky, M.	2014	Impact of prior treatment status and reasons for discontinuation on the efficacy and safety of fingolimod: Subgroup analyses of the Fingolimod Research Evaluating Effects of Daily Oral Therapy in Multiple Sclerosis (FREEDOMS) study	10.1016/j.msard.2013.10.006	Intervention
Lambe, J.	2021	Modulation of Retinal Atrophy With Rituximab in Multiple Sclerosis	10.1212/wnl.0000000000011933	Outcome
Lanzillo, R.	2020	Clinical predictors of Dimethyl Fumarate response in multiple sclerosis: a real life multicentre study	10.1016/j.msard.2019.101871	Study design
Lategan, T. W.	2021	Pharmacokinetics and Bioavailability of Monomethyl Fumarate Following a Single Oral Dose of Bafiertam TM (Monomethyl Fumarate) or Tecfidera R (Dimethyl Fumarate)	10.1007/s40263-021-00799-9	Population
Lefort, M.	2022	Disability progression in multiple sclerosis patients using early first-line treatments	10.1111/ene.15422	Outcome
Lily, O.	2006	Disease-specific quality of life in multiple sclerosis: the effect of disease modifying treatment	10.1177/1352458506070946	No confounder adjustment
Lopez-Caneda, C.	2023	Real-world persistence to first-line DMTs in relapsing-remitting multiple sclerosis	10.1016/j.msard.2023.104909	Outcome
Lorefice, L.	2021	The Dimethyl Fumarate Experience: A Handy Drug With Broad Clinical Utility	10.3389/fneur.2021.679355	Study design
Ludwig, M. D.	2016	Long-term treatment with low dose naltrexone maintains stable health in patients with multiple sclerosis	10.1177/2055217316672242	No confounder adjustment
Martins, P.	2023	Cost-effectiveness Analysis of Ocrelizumab for the Treatment of Relapsing and Primary Progressive Multiple Sclerosis in Portugal	10.1007/s41669-022-00381-z	Outcome
Masjedi, S. S.	2021	Assessment of fingolimod versus dimethyl fumarate for the treatment of multiple sclerosis; a 24-month follow-up study		Population
Masuda, H.	2020	Comparison of brain atrophy in patients with multiple sclerosis treated with first- versus second-generation disease modifying therapy without clinical relapse	10.1111/ene.14335	Outcome
Meca-Lallana, J. E.	2021	Four-year safety and effectiveness data from patients with multiple sclerosis treated with fingolimod: The Spanish GILENYA registry	10.1371/journal.pone.0258437	Intervention
Meca-Lallana, J. E.	2010	Effects of glatiramer acetate on spasticity in previously interferon-beta-treated and treatment-naive patients with relapsing-remitting multiple sclerosis: a prospective, nonrandomized, open-label, uncontrolled, observational pilot study	10.1016/j.clinthera.2010.06.005	Study design
Mesaros, S.	2008	Clinical and conventional MRI predictors of disability and brain atrophy accumulation in RRMS. A large scale, short-term follow-up study	10.1007/s00415-008-0924-5	No confounder adjustment
Mirabella, M.	2018	Safety and Efficacy of Dimethyl Fumarate in Multiple Sclerosis: An Italian, Multicenter, Real-World Study	10.1007/s40263-018-0543-3	No confounder adjustment
Naismith, R. T.	2020	Diroximel Fumarate Demonstrates an Improved Gastrointestinal Tolerability Profile Compared with Dimethyl Fumarate in Patients with Relapsing-Remitting Multiple Sclerosis: Results from the Randomized, Double-Blind, Phase III EVOLVE-MS-2 Study	10.1007/s40263-020-00700-0	Outcome
Naismith, R. T.	2020	Diroximel fumarate (DRF) in patients with relapsing-remitting multiple sclerosis: Interim safety and efficacy results from the phase 3 EVOLVE-MS-1 study	10.1177/1352458519881761	No confounder adjustment
Oreja-Guevara, C.	2009	[Glatiramer acetate in interferon beta non respondent relapsing-remitting multiple sclerosis]		Language
Papadopoulos, D.	2018	Oral Disease-Modifying Treatments for Relapsing Multiple Sclerosis: A Likelihood to Achieve No Evidence of Disease Activity or Harm Analysis	10.1007/s40263-018-0547-z	Study design
Patel, C.	2021	beta-endorphin and opioid growth factor as biomarkers of physical ability in multiple sclerosis	10.1016/j.msard.2021.102868	Study design
Pavelek, Z.	2020	Comparison of Therapies in MS Patients After the First Demyelinating Event in Real Clinical Practice in the Czech Republic: Data From the National Registry ReMuS	10.3389/fneur.2020.593527	No confounder adjustment
Pham Nguyen, T. P.	2021	Multiple sclerosis hospitalizations among users of oral disease-modifying therapies	10.1016/j.msard.2021.102944	Population

Prosperini, L.	2021	Prognostic Accuracy of NEDA-3 in Long-term Outcomes of Multiple Sclerosis	10.1212/nxi.0000000000001059	Outcome
Prosperini, L.	2020	Minimal evidence of disease activity (MEDA) in relapsing-remitting multiple sclerosis	10.1136/jnnp-2019-322348	Outcome
Putzki, N.	2010	Natalizumab reduces clinical and MRI activity in multiple sclerosis patients with high disease activity: results from a multicenter study in Switzerland	10.1159/000276400	No confounder adjustment
Puz, P.	2016	Safety and efficacy of fingolimod and natalizumab in multiple sclerosis after the failure of first-line therapy: Single center experience based on the treatment of forty-four patients	10.12659/msm.898270	No confounder adjustment
Repovic, P.	2021	Effectiveness of Dimethyl Fumarate in Patients With Relapsing Multiple Sclerosis Switching After Suboptimal Response to Glatiramer Acetate, Including Patients With Early Multiple Sclerosis: Subgroup Analysis of RESPOND	10.1007/s40120-020-00223-2	Study design
Reyes, S.	2020	Socioeconomic status and disease-modifying therapy prescribing patterns in people with multiple sclerosis	10.1016/j.msard.2020.102024	Study design
Rio, J.	2012	Change in the clinical activity of multiple sclerosis after treatment switch for suboptimal response	10.1111/j.1468-1331.2011.03648.x	No confounder adjustment
Rocha, R.	2022	Red cell distribution width as a predictor of disability in relapsing-remitting multiple sclerosis	10.1016/j.neurop.2022.02.001	Study design
Rodrigues, R.	2021	Therapeutic inertia in relapsing-remitting multiple sclerosis	10.1016/j.msard.2021.103176	Outcome
Romeo, M.	2013	Clinical and MRI predictors of response to interferon-beta and glatiramer acetate in relapsing-remitting multiple sclerosis patients	10.1111/ene.12119	No confounder adjustment
Rossi, S.	2013	Effect of glatiramer acetate on disease reactivation in MS patients discontinuing natalizumab	10.1111/j.1468-1331.2012.03794.x	No confounder adjustment
Sabin, J.	2020	Tolerability and safety of dimethyl fumarate in relapsing multiple sclerosis: a prospective observational multicenter study in a real-life Spanish population	10.1007/s00415-020-09848-7	Study design
Sacca, F.	2019	Determinants of therapy switch in multiple sclerosis treatment-naive patients: A real-life study	10.1177/1352458518790390	Outcome
Salter, A.	2020	Comorbidity is associated with disease activity in MS: Findings from the CombiRx trial	10.1212/wnl.0000000000010024	Study design
Simoneau, G.	2022	Overall and patient-level comparative effectiveness of dimethyl fumarate and fingolimod: A precision medicine application to the Observatoire Francais de la Sclerose en Plaques registry	10.1177/20552173221116591	No confounder adjustment
Singhal, T.	2017	The Effect of Glatiramer Acetate on Spinal Cord Volume in Relapsing-Remitting Multiple Sclerosis	10.1111/jon.12378	Outcome
Solmaz, I.	2022	Newer disease modifying treatments in pediatric onset multiple sclerosis: Experience from a single center	10.1016/j.ejpn.2022.06.013	Population
Souza, K. M.	2020	Effectiveness of first-line treatment for relapsing-remitting multiple sclerosis in Brazil: A 16-year non-concurrent cohort study	10.1371/journal.pone.0238476	Outcome
Spelman, T.	2023	Switching to natalizumab or fingolimod in multiple sclerosis: Comparative effectiveness and effect of pre-switch disease activity	10.1016/j.msard.2022.104477	Intervention
Stampanoni Bassi, M.	2020	Cerebrospinal fluid inflammatory biomarkers predicting interferon-beta response in MS patients	10.1177/1756286420970833	Population
Tanaka, E.	2023	Effect of smoking on disease activity in multiple sclerosis patients treated with dimethyl fumarate or fingolimod	10.1016/j.msard.2023.104513	Intervention
Topic, A.	2022	The Effects of Disease-Modifying Therapies on Oxidative Stress in Patients with Relapsing-Remitting Multiple Sclerosis	10.1097/wnf.0000000000000519	Outcome
Vollmer, B.	2019	Comparative discontinuation, effectiveness, and switching practices of dimethyl fumarate and fingolimod at 36-month follow-up	10.1016/j.jns.2019.116498	Population
Vollmer, B.	2018	Discontinuation and comparative effectiveness of dimethyl fumarate and fingolimod in 2 centers	10.1212/cpj.0000000000000487	Population
Vollmer, B. L.	2019	Natalizumab versus fingolimod and dimethyl fumarate in multiple sclerosis treatment	10.1002/acn3.700	Population
Vollmer, B. L.	2020	Rituximab versus natalizumab, fingolimod, and dimethyl fumarate in multiple sclerosis treatment	10.1002/acn3.51111	Population
Vollmer, B. L.	2022	Evolution of Disease Modifying Therapy Benefits and Risks: An Argument for De-escalation as a Treatment Paradigm for Patients With Multiple Sclerosis	10.3389/fneur.2021.799138	No confounder adjustment
Wawrzyniak, S.	2019	Early predictors of injectable disease modifying drugs suboptimal response based on clinical and radiological data assessment in Polish Multiple Sclerosis patients	10.5603/PJNNS.a2019.0010	Study design
Wundes, A.	2021	Improved gastrointestinal profile with diroximel fumarate is associated with a positive impact on quality of life compared with dimethyl fumarate: results from the randomized, double-blind, phase III EVOLVE-MS-2 study	10.1177/1756286421993999	Outcome
Wynn, D.	2020	Monomethyl fumarate has better gastrointestinal tolerability profile compared with dimethyl fumarate	10.1016/j.msard.2020.102335	Outcome
Ytterberg, C.	2007	Combination therapy with interferon-beta and glatiramer acetate in multiple sclerosis	10.1111/j.1600-0404.2007.00801.x	Population

Zanetta, C.	2023	Correction to: Effectiveness and safety profile of cladribine in an Italian real-life cohort of relapsing-remitting multiple sclerosis patients: a monocentric longitudinal observational study (Journal of Neurology, (2023), 270, 7, (3553-3564), 10.1007/s00415-023-11700-7)	10.1007/s00415-023-11772-5	Publication type
Zanghi, A.	2020	Pregnancy and the Postpartum Period in Women With Relapsing-Remitting Multiple Sclerosis Treated With Old and New Disease-Modifying Treatments: A Real-World Multicenter Experience	10.3389/fneur.2020.00105	Study design
Zanghi, A.	2021	First-line therapies in late-onset multiple sclerosis: An Italian registry study	10.1111/ene.15006	Outcome
Zanghi, A.	2022	Real world comparison of teriflunomide and dimethyl fumarate in naive relapsing multiple sclerosis patients: Evidence from the Italian MS register	10.1016/j.msard.2022.103489	Outcome
Zhu, W.	2022	Patterns of Utilization and Expenditure Across Multiple Sclerosis Disease-Modifying Therapies: A Retrospective Cohort Study Using Claims Data from a Commercially Insured Population in the United States, 2010-2019	10.1007/s40120-022-00358-4	Outcome
Ziemssen, T.	2022	Real-world patient characteristics, treatment patterns and costs in relapsing multiple sclerosis patients treated with glatiramer acetate, dimethyl fumarate or teriflunomide in Germany	10.2217/nmt-2021-0031	Outcome
Ziemssen, T.	2016	QualiCOP: real-world effectiveness, tolerability, and quality of life in patients with relapsing-remitting multiple sclerosis treated with glatiramer acetate, treatment-naive patients, and previously treated patients	10.1007/s00415-016-8058-7	No confounder adjustment

Anhang F Datensatz

Siehe separate Datei "Anhang F Datensatz.xlsx".

Anhang E Offenlegung von Beziehungen der externen Sachverständigen

Im Folgenden sind die Beziehungen der externen Sachverständigen zusammenfassend dargestellt. Alle Informationen beruhen auf Selbstangaben der einzelnen Personen anhand des „Formblatts zur Offenlegung von Beziehungen“. Die in diesem Formblatt aufgeführten Fragen finden sich im Anschluss an diese Zusammenfassung.

Externe Sachverständige

Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6	Frage 7
Lars G. Hemkens	ja	ja	ja	ja	nein	nein	ja
Julian Hirt	ja	nein	nein	ja	nein	nein	ja
Perrine Janiaud	ja	nein	nein	ja	nein	nein	ja
Tim Wölfle	ja	nein	nein	nein	ja	nein	ja

Im „Formblatt zur Offenlegung von Beziehungen“ wurden folgende 7 Fragen gestellt:

Frage 1: Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor bei einer Einrichtung des Gesundheitswesens (z. B. einer Klinik, einer Einrichtung der Selbstverwaltung, einer Fachgesellschaft, einem Auftragsforschungsinstitut), einem pharmazeutischen Unternehmen, einem Medizinproduktehersteller oder einem industriellen Interessenverband angestellt oder für diese / dieses / diesen selbstständig oder ehrenamtlich tätig bzw. sind oder waren Sie freiberuflich in eigener Praxis tätig?

Frage 2: Beraten Sie oder haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor eine Einrichtung des Gesundheitswesens (z. B. eine Klinik, eine Einrichtung der Selbstverwaltung, eine Fachgesellschaft, ein Auftragsforschungsinstitut), ein pharmazeutisches Unternehmen, einen Medizinproduktehersteller oder einen industriellen Interessenverband beraten (z. B. als Gutachter/-in, Sachverständige/r, in Zusammenhang mit klinischen Studien als Mitglied eines sogenannten Advisory Boards / eines Data Safety Monitoring Boards [DSMB] oder Steering Committees)?

Frage 3: Haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor direkt oder indirekt von einer Einrichtung des Gesundheitswesens (z. B. einer Klinik, einer Einrichtung der Selbstverwaltung, einer Fachgesellschaft, einem Auftragsforschungsinstitut), einem pharmazeutischen Unternehmen, einem Medizinproduktehersteller oder einem industriellen Interessenverband Honorare erhalten (z. B. für Vorträge, Schulungstätigkeiten, Stellungnahmen oder Artikel)?

Frage 4: Haben Sie oder hat Ihr Arbeitgeber bzw. Ihre Praxis oder die Institution, für die Sie ehrenamtlich tätig sind, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einer Einrichtung des Gesundheitswesens (z. B. einer Klinik, einer Einrichtung der Selbstverwaltung, einer Fachgesellschaft, einem Auftragsforschungsinstitut), einem pharmazeutischen Unternehmen, einem Medizinproduktehersteller oder einem industriellen Interessenverband sogenannte Drittmittel erhalten (d. h. finanzielle Unterstützung z. B. für Forschungsaktivitäten, die Durchführung klinischer Studien, andere wissenschaftliche Leistungen oder Patentanmeldungen)? Sofern Sie in einer größeren Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, z. B. Klinikabteilung, Forschungsgruppe.

Frage 5: Haben Sie oder hat Ihr Arbeitgeber bzw. Ihre Praxis oder die Institution, für die Sie ehrenamtlich tätig sind, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen, z. B. Ausrüstung, Personal, Unterstützung bei der Ausrichtung einer Veranstaltung, Übernahme von Reisekosten oder Teilnahmegebühren für Fortbildungen / Kongresse erhalten von einer Einrichtung des Gesundheitswesens (z. B. einer Klinik, einer Einrichtung der Selbstverwaltung, einer Fachgesellschaft, einem Auftragsforschungsinstitut), einem pharmazeutischen Unternehmen,

einem Medizinproduktehersteller oder einem industriellen Interessenverband? Sofern Sie in einer größeren Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, z. B. Klinikabteilung, Forschungsgruppe.

Frage 6: Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile einer Einrichtung des Gesundheitswesens (z. B. einer Klinik, einem Auftragsforschungsinstitut), eines pharmazeutischen Unternehmens, eines Medizinprodukteherstellers oder eines industriellen Interessenverbands? Besitzen Sie Anteile eines sogenannten Branchenfonds, der auf pharmazeutische Unternehmen oder Medizinproduktehersteller ausgerichtet ist? Besitzen Sie Patente für ein pharmazeutisches Erzeugnis, ein Medizinprodukt, eine medizinische Methode oder Gebrauchsmuster für ein pharmazeutisches Erzeugnis oder ein Medizinprodukt?

Frage 7: Sind oder waren Sie jemals an der Erstellung einer medizinischen Leitlinie oder klinischen Studie beteiligt, die eine mit diesem Projekt vergleichbare Thematik behandelt/e? Gibt es sonstige Umstände, die aus Sicht von unvoreingenommenen Betrachtenden als Interessenkonflikt bewertet werden können, z. B. Aktivitäten in gesundheitsbezogenen Interessengruppierungen bzw. Selbsthilfegruppen, politische, akademische, wissenschaftliche oder persönliche Interessen?