

Evaluation von Maßnahmen zur Aufwandsreduktion bei der systematischen Identifikation und der Erhebung von Confoundern

A decorative horizontal bar spanning the width of the page, composed of 18 small squares in various shades of blue and grey. The word 'ARBEITSPAPIER' is centered in white text on a dark blue rectangular background that covers the middle portion of this bar.

ARBEITSPAPIER

Projekt: GA25-02

Version: 1.0

Stand: 30.09.2026

IQWiG-Berichte – Nr. 2361

DOI: 10.60584/GA25-02

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Evaluation von Maßnahmen zur Aufwandsreduktion bei der systematischen Identifikation und der Erhebung von Confoundern

Auftraggeber

Bearbeitung im Rahmen des Generalauftrags

Interne Projektnummer

GA25-02

DOI-URL

<https://doi.org/10.60584/GA25-02>

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: info@iqwig.de
Internet: www.iqwig.de

ISSN: 1864-2500

Zitiervorschlag

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Evaluation von Maßnahmen zur Aufwandsreduktion bei der systematischen Identifikation und der Erhebung von Confoundern; Arbeitspapier [online]. 2026 [Zugriff: TT.MM.JJJJ]. URL: <https://doi.org/10.60584/GA25-02>.

Schlagwörter

Störfaktoren (Epidemiologie), Informationsspeicherung und -Retrieval, Nutzenbewertung, Konzept

Keywords

Confounding Factors – Epidemiologic, Information Storage and Retrieval, Benefit Assessment, Concept

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 und der externen Revierinnen und Reviewer zur Offenlegung von Beziehungen sind in Anhang I dargestellt. Es wurden keine Interessenkonflikte festgestellt, die die fachliche Unabhängigkeit im Hinblick auf eine Bearbeitung des vorliegenden Auftrags gefährden.

Externe Sachverständige

- Vanessa Didelez, Leibniz-Institut für Präventionsforschung und Epidemiologie – BIPS GmbH, Bremen

Externe Sachverständige für die Vor- und Nachbereitung und Moderation des Workshops

- Jessica L. Rohmann, QUEST Center for Responsible Research, Berlin Institute of Health at Charité – Universitätsmedizin Berlin

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

Mitarbeiterinnen und Mitarbeiter des IQWiG

- Jana Göbel
- Lars Beckmann
- Simone Johner
- Anne-Kathrin Petri
- Volker Vervölgyi
- Siw Waffenschmidt

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	viii
Abbildungsverzeichnis	x
Abkürzungsverzeichnis.....	xi
Glossar	xii
Kurzfassung	xv
1 Hintergrund.....	1
2 Fragestellung.....	2
3 Projektverlauf.....	3
4 Methoden	4
4.1 Aufwandsreduktion bei der Literaturrecherche für die systematische Confounderidentifikation	4
4.1.1 Confounderidentifikation über systematische Übersichten und die darin enthaltenen Primärstudien	5
4.1.2 Effizienz der Verfahren.....	8
4.1.2.1 Ergebnissicherheit und potenzieller Informationsverlust	8
4.1.2.2 Aufwandsreduktion	9
4.2 Aufwandsreduktion bei der Datenerhebung	9
4.2.1 Orientierende Informationsbeschaffung	9
4.2.2 Prüfung verschiedener Maßnahmen zur Aufwandsreduktion.....	10
4.3 Ableitung von Anforderungen an eine systematische Confounderidentifikation (Handlungsempfehlungen)	11
5 Ergebnisse	12
5.1 Aufwandsreduktion bei der Literaturrecherche.....	12
5.1.1 Confounderidentifikation über systematische Übersichten und die darin enthaltenen Primärstudien	12
5.1.1.1 Beispiel RRMS	12
5.1.1.1.1 Ergebnisse der Informationsbeschaffung	12
5.1.1.1.2 Ergebnisse der Auswertung	16
5.1.1.2 Beispiel MCL.....	25
5.1.1.2.1 Ergebnisse der Informationsbeschaffung	25
5.1.1.2.2 Ergebnisse der Auswertung	28
5.1.1.3 Zusammenfassung Confounderidentifikation über systematische Übersichten und die darin enthaltenen Primärstudien.....	31

5.1.2	Einfluss des Sättigungskriteriums auf die Identifikation potenzieller Confounder-Variablen	32
5.2	Aufwandsreduktion bei der Datenerhebung	35
5.2.1	Systematische Ansätze zur Reduktion eines Confounder-Sets	35
5.2.1.1	Anlegen von Schwellenwerten hinsichtlich der Häufigkeit der Nennung in der Literatur	37
5.2.1.2	Berücksichtigung der Art der Quelle	38
5.2.1.3	Berücksichtigung der Qualität der Quelle	40
5.2.1.4	Berücksichtigung von Korrelationen	42
5.2.1.5	Fokussierung auf Faktoren, die die Therapieentscheidung beeinflussen	44
5.2.1.6	Verzicht auf instrumentelle Variablen	45
5.2.1.7	Erarbeitung gerichteter azyklischer Graphen	46
5.2.1.7.1	Methodische Grundsätze	46
5.2.1.7.2	Workshop zur Erstellung eines gerichteten azyklischen Graphen	49
5.2.1.7.3	Ergebnis des DAG-Workshops	52
5.2.1.7.4	Zusammenfassung	59
5.2.1.8	Relevanz der Confounder – Bewertung der Stärke des Einflusses auf die Effektschätzung	60
5.2.2	Zusammenfassende Gegenüberstellung geeigneter und nicht geeigneter Maßnahmen zur Reduktion eines Confounder-Sets	64
5.3	Vorschlag für die Anforderung an eine systematische Confounderidentifikation und -erhebung	66
5.3.1	Schritt 1: Identifikation	68
5.3.2	Schritt 2: Zusammenführung	70
5.3.3	Schritt 3: Reduktion	70
6	Diskussion	75
7	Fazit	79
8	Literatur	81
Anhang A	Suchstrategien	91
A.1	Recherchen zur Confounderidentifikation über systematische Übersichten und die darin enthaltenen Primärstudien	91
A.1.1	RRMS	91
A.1.1.1	Suche nach systematischen Übersichten	91
A.1.1.2	Suche nach Primärstudien (zur Aufwandsschätzung)	92
A.1.2	MCL	93
A.1.2.1	Suche nach systematischen Übersichten	93

Anhang B	Bewertung der Qualität der berücksichtigten systematischen Übersichten	95
B.1	Beispiel RRMS	95
B.2	Beispiel MCL	96
Anhang C	Studienlisten	97
C.1	Beispiel RRMS	97
C.1.1	Liste der relevanten systematischen Übersichten	97
C.1.2	Liste der im Volltext ausgeschlossenen systematischen Übersichten	97
C.1.3	Liste der eingeschlossenen Primärpublikationen identifiziert durch systematische Übersichten	98
C.1.4	Liste der im Volltext ausgeschlossenen Primärpublikationen identifiziert durch systematische Übersichten	101
C.1.5	Liste der eingeschlossenen Primärpublikationen aus dem Datensatz der externen Sachverständigen	102
C.2	Beispiel MCL	105
C.2.1	Liste der relevanten systematischen Übersichten	105
C.2.2	Liste der im Volltext ausgeschlossenen systematischen Übersichten	105
C.2.3	Liste der eingeschlossenen Primärpublikationen identifiziert durch systematische Übersichten	107
C.2.4	Liste der im Volltext ausgeschlossenen Primärpublikationen identifiziert durch systematische Übersichten	109
Anhang D	Anwendung eines Sättigungskriteriums	111
Anhang E	Erarbeitung eines gerichteten azyklischen Graphen	112
E.1	Zeitliches Raster	112
E.2	Matrix der Zusammenhänge und DAGitty Model Code	113
E.2.1	Matrix zur Dokumentation der Zusammenhänge zwischen den Variablen	113
E.2.2	DAGitty Modell-Code des finalen DAG (Anordnung im zeitlichen Raster)	114
E.3	Tabellarische Darstellung der weggelassenen Pfeile zwischen einzelnen Confoundern	118
E.4	Dokumentation der im Workshop vorgenommenen Änderungen am Confounder-Set aus GA23-02	121
E.5	Dokumentation der im Nachgang zum Workshop vorgenommenen Anpassungen der Zusammenhänge	123
Anhang F	Online-Umfrage zur Bewertung der Relevanz der Confounder	124
F.1	IQWiG-Begleitschreiben zur Online-Umfrage	124
F.2	Online-Umfrage	126
F.3	Ergebnis der Online-Umfrage	133

Anhang G	Vorschlag für ein Formblatt zur Offenlegung von Beziehungen bei der Einbindung von Expertinnen und Experten in die Confounderidentifikation	135
Anhang H	Handlungsempfehlungen – Checkliste	137
Anhang I	Offenlegung von Beziehungen der externen Sachverständigen	140

Tabellenverzeichnis

	Seite
Tabelle 1: Ein- und Ausschlusskriterien für das Literaturscreening.....	6
Tabelle 2: Berücksichtigte systematische Übersichten (Beispiel RRMS)	14
Tabelle 3: Abgleich mit Referenzstandard (Beispiel RRMS).....	17
Tabelle 4: Anzahl der identifizierten potenziellen Confounder aus dem Referenzset (N = 26) in Abhängigkeit des Datensatzes (Beispiel RRMS).....	18
Tabelle 5: Anzahl der identifizierten potenziellen Confounder aus dem Referenzset (N = 26) in Abhängigkeit des Studientyps bzw. der Extraktionsquelle (Beispiel RRMS) ...	19
Tabelle 6: Anzahl der identifizierten potenziellen Confounder aus dem Referenzset (N = 26) in Abhängigkeit der Aktualität	20
Tabelle 7: Identifizierte potenzielle Confounder ausgewertet nach erstmaliger Nennung	22
Tabelle 8: Vergleich der Aufwände bei verschiedenen Vorgehensweisen der Informationsbeschaffung und -auswertung (Beispiel RRMS)	24
Tabelle 9: Berücksichtigte systematische Übersichten.....	27
Tabelle 10: Abgleich mit Referenzstandard (Beispiel MCL)	29
Tabelle 11: Anzahl der identifizierten potenziellen Confounder in Abhängigkeit des Studientyps bzw. der Extraktionsquelle (Beispiel MCL)	30
Tabelle 12: Anzahl der identifizierten potenziellen Confounder in Abhängigkeit der Aktualität (Beispiel MCL)	30
Tabelle 13: Sättigungskriterium – Ergebnis der Simulation einer zufälligen Anordnung (10 000-mal) der 80 Publikationen bei Anwendung eines Sättigungskriteriums von 15, 10 oder 5.....	33
Tabelle 14: Sättigungskriterium – Ergebnis der Simulation einer zufälligen Anordnung (10 000-mal) der 27 RCTs bei Anwendung eines Sättigungskriteriums von 15, 10 oder 5.....	35
Tabelle 15: Schritte der Planung und Durchführung des Workshops zur Erstellung eines DAG	49
Tabelle 16: Teilnehmende klinische / methodische Expertinnen und Experten des Workshops.....	50
Tabelle 17: Dokumentation der weggelassenen Pfeile zur Therapieentscheidung und zu den Endpunkten	53
Tabelle 18: Darstellung der 3 identifizierten minimal ausreichenden Adjustierungssets.....	57
Tabelle 19: Bewertung der Informationsbeschaffung zur systematischen Übersicht Prosperini 2023.....	95
Tabelle 20: Bewertung der Informationsbeschaffung zur systematischen Übersicht Gonzalez-Lorenzo 2024	95

Tabelle 21: Bewertung der Informationsbeschaffung zur systematischen Übersicht Tramacere 2023.....	95
Tabelle 22: Bewertung der Informationsbeschaffung zur systematischen Übersicht Sancho 2024.....	96
Tabelle 23: Bewertung der Informationsbeschaffung zur systematischen Übersicht National Institute for Health and Care Excellence (NICE) 2021	96
Tabelle 24: Bewertung der Informationsbeschaffung zur systematischen Übersicht Wu 2024	96
Tabelle 25: Sättigungskriterium – Ergebnis der zeitlichen Anordnung der 80 Publikationen bei Anwendung eines Sättigungskriteriums von 15, 10 oder 5	111
Tabelle 26: Dokumentation weggelassener Pfeile zwischen einzelnen Confoundern	118
Tabelle 27: Identifiziertes Set potenzieller Confounder aus GA23-02 inklusive vorgeschlagener Operationalisierung in der Indikation RRMS und den im Workshop vorgenommenen Änderungen durch die klinischen Expertinnen und Experten.....	121
Tabelle 28: Änderung der besprochenen Zusammenhänge zwischen den einzelnen Confoundern im Nachgang des Workshops	123
Tabelle 29: Ergebnis der Online-Umfrage (n = 18) zur Bewertung der Stärke des Einflusses aller einzelnen Confounder auf die Therapieentscheidung bzw. auf die Endpunkte.....	133

Abbildungsverzeichnis

	Seite
Abbildung 1: Ergebnis der Informationsbeschaffung und der Studienselektion nach systematischen Übersichten (Beispiel RRMS)	13
Abbildung 2: Ergebnis der Sichtung der eingeschlossenen Publikationen (Beispiel RRMS)....	15
Abbildung 3: Ergebnis der Informationsbeschaffung und der Studienselektion nach systematischen Übersichten (Beispiel MCL)	26
Abbildung 4: Ergebnis der Sichtung der eingeschlossenen Publikationen (Beispiel MCL)	28
Abbildung 5: Beispielhafte Konstellationen 2 korrelierter Confounder (1 und 2).....	43
Abbildung 6: Darstellung der verschiedenen Variablentypen und ihren jeweiligen Zusammenhängen zu Therapieentscheidung und Endpunkt in einem gerichteten azyklischen Graphen	47
Abbildung 7: Finaler gerichteter azyklischer Graph (DAG)	55
Abbildung 8: Boxplots der Ergebnisse der Online-Umfrage	63
Abbildung 9: Erweiterung des in GA23-02 vorgeschlagenen Verfahrens der systematischen Confounderidentifikation um Schritt 3.....	67
Abbildung 10: Anordnung der Confounder-Variablen im zeitlichen Raster (Ausgangspunkt für den DAG-Workshop)	112
Abbildung 11: Dokumentation der Zusammenhänge zwischen Therapieentscheidung, Endpunkt und den einzelnen Variablen	113

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AbD	anwendungsbegleitende Datenerhebung
DAG	Directed acyclic Graph (gerichteter azyklischer Graph)
DMSG	Deutsche Multiple Sklerose Gesellschaft
EMA	European Medicines Agency
FDA	Food and Drug Administration
GBA	Gemeinsamer Bundesausschuss
HTA	Health Technology Assessment
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
MCL	Mantle Cell Lymphoma (Mantelzell-Lymphom)
MS	multiple Sklerose
NICE	National Institute for Health and Care Excellence
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
RRMS	Relapsing-remitting multiple Sclerosis (schubförmig remittierende multiple Sklerose)

Glossar

Für weitergehende Literatur zu den im Folgenden beschriebenen Begriffsdefinitionen eignen sich z. B. die Publikationen Williams 2018 [1], Catalogue of bias collaboration [2], Sendor 2022 [3], Rhodes 1999 [4], Hernan 2022 [5], Hernan und Robins 2024 [6] und Bashir 2026 [7].

Confounder

Ein Confounder ist eine Variable, die sowohl die Therapieentscheidung als auch den Endpunkt beeinflusst und somit einen Behandlungseffekt verfälschen kann, wenn der Einfluss nicht berücksichtigt wird. In nicht randomisierten vergleichenden Studien soll durch die Adjustierung für Confoundern die fehlende Strukturgleichheit zwischen den Behandlungsgruppen ausgeglichen werden. In der statistischen Analyse kann jedoch nur für bekannte und gemessene Confounder adjustiert werden, die Verzerrung durch unbekannte oder nicht gemessene bzw. nicht messbare Confounder bleibt bestehen (sog. residuelles Confounding – siehe unten).

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.

Collider

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

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.

Minimal ausreichendes Adjustierungsset (minimal sufficient adjustment set)

Ein minimal ausreichendes Adjustierungsset kann mit Hilfe eines DAG ermittelt werden und beschreibt die kleinst mögliche Liste bzw. Teilmenge an identifizierten Confoundern, die nötig ist, um alle identifizierten Confounder adäquat in der Auswertung zu berücksichtigen. Es werden so wenige Confounder wie möglich und so viele wie notwendig berücksichtigt. In der Theorie ist das Ziel des DAG ein (im zugrundeliegenden kausalen Modell) unverzerrtes Ergebnis zu erhalten. Ein unverzerrtes Ergebnis könnte jedoch nur unter der in der Praxis sehr unwahrscheinlichen Voraussetzung erzielt werden, dass keine unbekannten Confounder oder residuelles Confounding vorliegen. Durch die Adjustierung für das minimal ausreichende Adjustierungsset werden alle „Hintertürpfade“ zwischen der Therapieentscheidung und Endpunkt geschlossen.

Instrumentenvariable

Eine Instrumentenvariable muss die folgenden 3 Bedingungen erfüllen: (1) die Variable ist mit der Intervention assoziiert, (2) hat nur über die Intervention Einfluss auf den Endpunkt und (3) die Variable und der Endpunkt haben keine gemeinsamen Ursachen.

Mediator

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

Multikollinearität

Multikollinearität in der Statistik bedeutet, dass 2 oder mehr unabhängige Variablen in einem Regressionsmodell stark miteinander korreliert sind. Dadurch wird die genaue Berechnung und Interpretation der einzelnen Variableneffekte erheblich erschwert.

Selection Bias

Ein Selection Bias entsteht, wenn Patientinnen und Patienten, ein Teil der Beobachtungsdauer (Follow-up) oder Ereignisse von der Analyse ausgeschlossen werden und die Selektion sowohl mit der Exposition als auch dem Endpunkt assoziiert ist.

Target-Trial-Emulation

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

Kurzfassung

Im Rahmen des Generalauftrags wurde das Thema Evaluation von Maßnahmen zur Aufwandsreduktion bei der systematischen Identifikation und der Erhebung von Confoundern bearbeitet.

Fragestellung

Ziele der vorliegenden Untersuchung sind aufbauend auf dem Arbeitspapier GA23-02:

- Prüfung von Möglichkeiten zur Aufwandsreduktion bei der systematischen Literaturrecherche für die Confounderidentifikation,
- Prüfung weiterer Möglichkeiten zur Aufwandsreduktion bei der nachfolgenden Datenerhebung (z. B. in einem Register) durch die Reduktion der Zahl der für die nicht randomisierte Studie zu erhebenden Confounder,
- Ableitung von Handlungsempfehlungen zur weiteren Aufwandsreduktion bei der systematischen Identifikation und der Erhebung von Confoundern im Sinne einer Weiterentwicklung der Handlungsempfehlungen aus dem Arbeitspapier GA23-02,

jeweils unter Berücksichtigung des potenziell damit einhergehenden Verlustes der Ergebnissicherheit.

Methoden

Aufwandsreduktion bei der Literaturrecherche für die systematische Confounderidentifikation

Im Rahmen dieses Projektes wurde anhand von 2 Anwendungsbeispielen (schubförmig remittierende multiple Sklerose [RRMS] und Mantelzell-Lymphom [MCL]) geprüft, ob ein alleiniger Rückgriff auf Sekundärliteratur (d. h. systematische Übersichten) zur Identifikation der Primärpublikationen und somit zur Confounderidentifikation genutzt werden kann und ob dies einen ausreichend sensitiven Ansatz darstellt. Hierzu wurden zunächst relevante systematische Übersichtsarbeiten identifiziert und anschließend basierend auf der Qualität der Informationsbeschaffung, Abdeckung bez. Fragestellung, Studientypen und Aktualität eine Auswahl getroffen. Die in den Übersichten eingeschlossenen Primärpublikationen wurden erfasst und die Primärpublikationen, die die Ein- und Ausschlusskriterien erfüllten, zur Extraktion der potenziellen Confounder-Variablen herangezogen. Die extrahierten Variablen wurden jeweils mit einem Referenzstandard abgeglichen. Der Referenzstandard war jeweils ein definiertes Set an Confoundern zum einen aus dem Projekt GA23-02 und zum anderen aus dem Beschluss zur Forderung einer AbD für den Wirkstoff Brexucabtagen-Autoleucel. Der potenzielle Informationsverlust wurde dargestellt. Darüber hinaus wurde anhand der

Ergebnisse der Einfluss des Studientyps, die Bedeutung der Aktualität der Recherche nach Primärstudien sowie der Einfluss eines Sättigungskriteriums geprüft. Abschließend wurde der Aufwand für die Studienidentifikation mittels bibliografischer Recherche nach Primärstudien mit der Studienidentifikation auf der Grundlage von systematischen Übersichten verglichen.

Aufwandsreduktion bei der Datenerhebung

Anhand einer orientierenden Recherche in PubMed wurde geprüft, ob in der Literatur Ansätze zur systematischen Reduktion eines Confounder-Sets basierend auf inhaltlichen Überlegungen unabhängig vom auszuwertenden Datensatz beschrieben werden.

Aufbauend auf den Untersuchungen in GA23-02 wurde geprüft, welche weiteren Möglichkeiten zur Aufwandsreduktion der Confoundererhebung bestehen. Dies wurde zum einen empirisch über den umfangreichen Datensatz zu den identifizierten potenziellen Confounder-Variablen in der Indikation RRMS untersucht. Zum anderen wurden die in der Anhörung zum vorläufigen Rapid Report A25-13 vorgebrachten Anregungen dahin gehend geprüft, ob sie weitere relevante Maßnahmen zur Aufwandsreduktion darstellen. Die als geeignet identifizierten Maßnahmen wurden auf die in GA23-02 identifizierten 28 Confounder angewendet und anhand dieses Beispiels diskutiert, welche Auswirkungen des Vorgehens auf die Ergebnissicherheit zu erwarten sind.

Ableitung von Anforderungen an eine systematische Confounderidentifikation (Handlungsempfehlungen)

Abschließend wurden anhand der gesichteten Literatur und den Ergebnissen der Untersuchungen konkrete Handlungsempfehlungen zur systematischen Identifikation und der Erhebung von Confoundern für nicht randomisierte vergleichende Studien aus GA23-02 im Sinne der Aufwandsreduktion weiterentwickelt. Hierbei wurde im Speziellen geprüft, inwieweit sich für Fragestellungen mit erwartbar weniger umfassender Studienlage (wie beispielsweise Fragestellungen der anwendungsbegleitenden Datenerhebung [AbD]) besondere Anforderungen ergeben.

Ergebnisse

Aufwandsreduktion bei der Literaturrecherche

Aus dem Beispiel MCL kann abgeleitet werden, dass bei Vorliegen aktueller, qualitativ hochwertiger systematischer Übersichten, die die gesamte Fragestellung abdecken, über die darin eingeschlossenen Studien ein Großteil der potenziell relevanten Confounder identifiziert werden kann.

Beim Beispiel RRMS konnte über die systematischen Übersichten nicht die gesamte Fragestellung abgedeckt werden und es wurde eine zusätzliche Suche nach Primärstudien für

einen Teilaspekt (Kohortenstudien für Glatirameracetat) notwendig. Es wurde deutlich, dass in dieser Situation keine Aufwandsreduktion möglich ist.

Beide Beispiele zeigen, dass der Informationsverlust am geringsten war, wenn alle verfügbaren Studientypen bzw. Quellen genutzt werden. Am Beispiel RRMS waren dies RCTs und Kohortenstudien; beim Beispiel MCL auch 1-armige Studien, da dort nur wenige andere Studientypen vorlagen.

Bei beiden Beispielen war das Heranziehen von Primärstudien mit einem Erscheinungsjahr bis 1 oder 2 Jahre vor der Erstellung des Referenzstandards notwendig, um nahezu alle potenziellen Confounder identifizieren zu können. Aus diesem Grund sollte die Recherche bei einem prospektiven Studiendesign zu Beginn der Datenerhebung – bzw. bei einem retrospektiven Studiendesign zum Zeitpunkt der Finalisierung des Studienprotokolls – nicht älter als 12 Monate sein.

Es konnte anhand des RRMS-Beispiels gezeigt werden, dass die Informationsbeschaffung über systematische Übersichten im Vergleich zum Vorgehen der externen Sachverständigen grundsätzlich zeitsparender und eine Möglichkeit zur deutlichen Aufwandsreduktion sein kann. Hier ist jedoch zu beachten, dass dies nur in Situationen möglich ist, in denen über systematische Übersichten die gesamte Fragestellung abgedeckt werden kann.

Aufwandsreduktion bei der Datenerhebung

Systematische Ansätze zur Reduktion eines Confounder-Sets

Durch die orientierende Recherche nach Publikationen zu systematischen Ansätzen zur Reduktion eines Confounder-Sets wurden 3 Publikationen identifiziert, die zumindest in Teilen dieses Thema behandeln. In diesen wurden prognostische Faktoren für den Endpunkt durch eingebundene klinische Expertinnen und Experten entsprechend ihrer Bedeutung kategorisiert und dadurch die wichtigsten Faktoren identifiziert. Diese Publikationen wurden im Abschnitt zur Relevanzbewertung der Confounder berücksichtigt. Weitere Ansätze zur Selektion von Confoundern, welche über die in der Anhörung bereits genannten Anregungen hinausgehen und nicht auf statistischen Verfahren basieren, wurden nicht gefunden.

Welche weiteren Möglichkeiten zur Aufwandsreduktion der Confoundererhebung bestehen, wurde zum einen empirisch über den umfangreichen Datensatz in der Indikation RRMS und zum anderen anhand der in der Anhörung zum vorläufigen Rapid Report A25-13 vorgebrachten Anregungen geprüft. Basierend auf dem RRMS-Datensatz konnten keine weitergehenden Kriterien, als die in der Anhörung bereits genannten, identifiziert werden.

Die folgenden Maßnahmen wurden als nicht zur systematischen Reduktion des Confounder-Sets geeignet eingeschätzt und werden daher in der Kurzfassung nicht näher beschrieben: Anlegen von Schwellenwerten hinsichtlich der Häufigkeit der Nennung in der Literatur,

Berücksichtigung der Art der Quelle sowie der Qualität der Quelle, Berücksichtigung von Korrelationen, Fokussierung auf Faktoren, die die Therapieentscheidung beeinflussen sowie der Verzicht auf instrumentelle Variablen.

Grundsätzlich ist zu beachten, dass einzelne Maßnahmen ggf. mit einem höheren Aufwand während der Planung der Studie einhergehen können, diese anschließend aber – insbesondere in einem prospektiven Studiendesign – den Aufwand der Erhebung, bspw. durch die Reduktion der Anzahl der im Register zu erhebenden Variablen, deutlich erleichtern. Insgesamt wird der Gesamtaufwand, bestehend aus Planung und Erhebung, betrachtet. Auch wenn bei den folgenden Ausführungen primär von einem prospektiven Studiendesign ausgegangen wird, gelten die daraus abgeleiteten Empfehlungen nichtsdestotrotz gleichermaßen für retrospektive Studien.

Erarbeitung gerichteter azyklischer Graphen

In mehreren Stellungnahmen wurde auf gerichtete azyklische Graphen (directed acyclic graphs [DAGs]) als Möglichkeit zur Identifikation eines Minimalsets an Confoundern und damit zur systematischen Reduktion des zu erhebenden Confounder-Sets hingewiesen.

Es wurde ein (2-teiliger) insgesamt 7,5-stündiger Workshop mit 4 praktizierenden Klinikerinnen und Klinikern durchgeführt. Zusätzlich zu den klinischen Expertinnen und Experten wurde eine Methodikerin mit Expertise in DAGs als externe Sachverständige eingebunden.

Das in GA23-02 identifizierte Confounder-Set umfasste insgesamt 28 Variablen. Die im Workshop eingebrachten begründeten Änderungen der Klinikerinnen und Kliniker führte zu einem Confounder-Set aus 31 Variablen. Mittels des DAG wurden 3 minimal ausreichende Adjustierungssets identifiziert, welche jeweils 27 Confounder umfassen. Jedes Confounder-Set für sich wäre theoretisch für eine vollständige Adjustierung ausreichend. Da jedoch in allen Sets die beiden nicht messbaren Confounder Erwartung an die Behandlung und Patientenwunsch enthalten sind, ist auch nach der Adjustierung für alle erhebbaren Confounder davon auszugehen, dass residuelles Confounding vorliegen wird.

Insgesamt wurde in dem Workshop unter Experteneinbindung ein DAG bestehend aus 31 Variablen für die vorliegende Fragestellung erstellt. Der Aufwand zur Erarbeitung des DAG wird als überschaubar eingeschätzt. Im vorliegenden Beispiel war mit 4 Confoundern zwar lediglich eine geringe Aufwandsreduktion möglich, allerdings besteht das Potenzial je nach Fragestellung bzw. Indikation eine höhere Anzahl an Confoundern einzusparen. Ein wesentlicher Vorteil der Erstellung eines DAG ist die transparente Dokumentation der zugrunde liegenden Annahmen. Unsicherheiten bezüglich getroffener Annahmen können gekennzeichnet und deren potenzielle Auswirkung auf das minimal ausreichende Adjustierungsset überprüft werden.

Relevanz der Confounder – Bewertung der Stärke des Einflusses auf die Effektschätzung

Das Ziel der zuvor beschriebenen DAG-Erstellung ist, diejenige Teilmenge der identifizierten Confounder zu ermitteln, für die mindestens adjustiert werden muss, um alle identifizierten Confounder adäquat in der Auswertung zu berücksichtigen. Die in diesem Schritt in den minimal ausreichenden Adjustierungssets identifizierten Confounder können jedoch einen deutlich unterschiedlichen Einfluss auf die Effektschätzung haben. Damit ist ein weiterer inhaltlich begründeter Ansatz ein Confounder-Set zu reduzieren, nur diejenigen Confounder auszuwählen, die den größtmöglichen Einfluss auf die Effektschätzung haben bzw. diejenigen aus dem Set zu entfernen, die den geringsten Einfluss aufweisen. Ein möglicher Ansatz sich dem Einfluss der Confounder auf die Effektschätzung anzunähern, ist, die inhaltliche Einschätzung der Relevanz der identifizierten Confounder zu ermitteln.

Um zu prüfen, wie eine Relevanzbewertung zur Reduktion von Confoundern beitragen kann, wurde im vorliegenden Projekt – analog zum Vorgehen der externen Sachverständigen in GA23-02 und den 3 identifizierten Publikationen – eine Umfrage zur Bewertung der Stärke des Einflusses der identifizierten Confounder durchgeführt. Die Bewertung erfolgte durch klinische Expertinnen und Experten mit Expertise in der entsprechenden Indikation (hier RRMS).

Ausgangspunkt für die Umfrage bildete das Ergebnis des DAG-Workshops (31 Confounder). Für jeden Confounder wurde auf einer 5-Punkt-Skala von 1 („sehr gering“) bis 5 („sehr stark“) die Stärke des Einflusses auf die Therapieentscheidung bzw. auf die Endpunkte abgefragt.

Insgesamt nahmen 18 Klinikerinnen und Klinikern an der Umfrage teil. Alle Teilnehmenden füllten die Umfrage vollständig aus. Um ein Ranking der Confounder zu ermöglichen, ist es notwendig die beiden getrennten Bewertungen (Stärke des Einflusses auf die Therapieentscheidung bzw. Endpunkte) zusammenzufassen. Hierzu gibt es verschiedene Möglichkeiten. Im Sinne eines pragmatischen Vorgehens wurde für jeden der Teilnehmenden und für jeden Confounder das arithmetische Mittel gewählt. Dies entspricht einer 1:1 Gewichtung. Das Ranking der Confounder wurde anhand des Medians vorgenommen.

Insgesamt wurde die Stärke des Einflusses von 10 der 31 Variablen durch die Klinikerinnen und Kliniker im Median mit ≥ 4 Punkten (d. h. stark oder sehr stark) bewertet. 3 Variablen wurden im Median mit ≤ 2 Punkten (d. h. gering oder sehr gering) bewertet.

Sollte eine Reduktion des über die vorherigen Schritte identifizierten Confounder-Sets notwendig sein, kann das aus der Umfrage abgeleitete Ranking der Confounder als Grundlage dienen. Hier sollte ein Cut-off gewählt werden, der sich an der erwarteten Bedeutung der Confounder orientiert. Dadurch werden diejenigen Confounder aus dem finalen Set entfernt, denen der geringste Einfluss beigemessen wird. Durch das Anlegen eines Cut-offs von ≤ 2 Punkten im Median werden diejenigen Confounder aus dem finalen Set entfernt, deren Stärke

des Einflusses als sehr gering (1 Punkt) oder gering (2 Punkte) bewertet wurde. Dies führt im vorliegenden Confounder-Set zu einer Reduktion von 3 Confoundern (Abstammung, Bildungsgrad und Art der beruflichen Tätigkeit).

Zusammenfassende Gegenüberstellung geeigneter und nicht geeigneter Maßnahmen zur Reduktion eines Confounder-Sets

Das Ziel der Reduktion eines Confounder-Sets muss immer im Kontext mit dem damit einhergehenden Verlust an Ergebnissicherheit betrachtet werden. Ein Ausschluss von Confoundern sollte deshalb immer inhaltlich begründet sein. Die Maßnahmen (Schwellenwerte, Berücksichtigung der Art oder Qualität der Quelle, Korrelationen, Fokussierung auf Faktoren die die Therapieentscheidung beeinflussen, Verzicht auf instrumentelle Variablen) haben sich als nicht zielführend für die Reduktion des Confounder-Sets erwiesen, da diese nicht ausreichend inhaltlich begründet sind.

Einzig die Erarbeitung eines DAG und ggf. die anschließende Bewertung der Confounder hinsichtlich der Stärke des Einflusses auf die Therapieentscheidung und die Endpunkte (und damit der vermuteten Stärke des Einflusses auf die Effektschätzung) stellen geeignete Maßnahmen zur systematischen und transparenten Reduktion eines Confounder-Sets dar. In einem DAG werden die der Fragestellung zugrunde liegenden Annahmen über die qualitativen kausalen Zusammenhänge explizit und damit transparent gemacht und die Reduktion basiert auf inhaltlichen Überlegungen. Auch wenn das DAG im vorliegenden Beispiel nicht zu einer relevanten Reduktion an zu erhebenden Confoundern geführt hat, ist es mindestens ein wichtiger Schritt zur Validierung des identifizierten Confounder-Sets und wird entsprechend als Maßnahme empfohlen. In anderen Fragestellungen kann mithilfe eines DAG ggf. eine relevante Reduktion an Confoundern erzielt werden.

Bei der DAG-Erstellung und der Bestimmung minimaler Adjustierungssets wird entsprechend der Methodik nicht berücksichtigt, wie stark ein Confounder auf die Therapieentscheidung bzw. den Endpunkt wirkt. Die Berücksichtigung dieser Stärke des Einflusses kann im letzten Schritt eine weitere Maßnahme zur Reduktion des Confounder-Sets darstellen. So können auf Basis der Einschätzung von klinischen Expertinnen und Experten die Confounder aus dem finalen Confounder-Set entfernt werden, denen nur ein untergeordneter Einfluss beigemessen wird. Entscheidend ist hier die Wahl des Cut-offs, der zur Reduktion des Confounder-Sets genutzt wird. Dabei ist es vertretbar, Confounder, denen nur eine sehr geringe und geringe Bedeutung beigemessen wird, aus dem finalen Confounder-Set zu entfernen. Wird die Mehrzahl an Confoundern als wichtig eingeordnet, ist keine relevante Aufwandsreduktion möglich, ohne das Risiko eines nicht abschätzbaren Verlusts an Ergebnissicherheit einzugehen. Im vorliegenden Beispiel hat die Anwendung dieses letzten Schritts ebenfalls nur zu einer sehr geringen Aufwandsreduktion hinsichtlich der Datenerhebung geführt.

Aufgrund des potenziell nur geringen Einsparpotenzials und der zusätzlichen Unsicherheiten, die dieser Schritt mit sich bringt, sollte er nur in Abhängigkeit der Anzahl an Confoundern und der für die Studie zur Verfügung stehenden Patientenzahl erfolgen. Der Schritt der Reduktion ist in Situationen nicht notwendig, in denen eine ausreichende Anzahl an Patientinnen und Patienten für die Studie zur Verfügung steht und die Anzahl identifizierter Confounder für die der durchzuführenden Studie zugrunde liegenden Datenquelle (z. B. ein klinisches Register) vertretbar ist. Letzteres ist insbesondere davon abhängig, wie viele Confounder bereits routinemäßig in der gewählten Datenquelle erhoben werden und wie viele Confounder zusätzlich für die spezifische Fragestellung in die Datenquelle aufgenommen werden müssten.

Vorschlag für die Anforderung an eine systematische Confounderidentifikation und -erhebung

Zur Evaluation von Interventionseffekten auf Basis eines nicht randomisierten Vergleichs, sollte im Rahmen der Studienplanung a priori ein Set potenzieller Confounder identifiziert werden. Hierfür wurden in GA23-02 Vorschläge für die konkreten Anforderungen an die systematische Confounderidentifikation erarbeitet (bestehend aus den 3 Schritten: Identifikation, Zusammenführung und Relevanzbewertung). Das Verfahren der systematischen Confounderidentifikation ist trotz der in GA23-02 beschriebenen Möglichkeiten zur Aufwandseinsparung ressourcenintensiv. Daher sollte im vorliegenden Arbeitspapier geprüft werden, welche weiteren Möglichkeiten zur Aufwandsreduktion bestehen, sowohl für die Confounderidentifikation als auch die zugehörige Erhebung der potenziellen Confounder. Aus diesen Gründen wurde das Verfahren im vorliegenden Projekt mit dem Ziel der Aufwandsreduktion weiterentwickelt.

Nach den in GA23-02 beschriebenen Schritten 1 zur Identifikation und 2 zur Zusammenführung sollte in einem 3. Schritt versucht werden, den Aufwand für die Confounder-Erhebung zu reduzieren. Hierzu sollten in einem 1. Teilschritt die angenommenen kausalen Zusammenhänge in einem DAG visualisiert werden. Dieser Schritt dient insbesondere der Identifikation des minimal ausreichenden Adjustierungssets. Gleichzeitig werden durch diesen Schritt die über die Literaturrecherche und die Interviews mit Klinikerinnen und Klinikern identifizierten und anschließend verdichteten Confounder-Sets validiert. Anschließend kann die Bewertung der Relevanz der Confounder durch klinische Expertinnen und Experten in einem 2. optionalen Teilschritt zur weiteren Reduktion des Confounder-Sets einbezogen werden.

Schritt 1: Identifikation

Zusätzlich zu den in GA23-02 beschriebenen Handlungsempfehlungen lassen sich aus den vorliegenden Untersuchungen die folgenden Empfehlungen ableiten: Im Sinne der Aufwandsreduktion können die in hochwertigen systematischen Übersichten enthaltenen Primärpublikationen für die Extraktion potenzieller Confounder-Variablen verwendet werden,

sofern durch die systematischen Übersichten die gesamte Fragestellung abgedeckt wird. Der Recherchezeitpunkt nach systematischen Übersichten bzw. Primärstudien sollte dabei nicht älter als 12 Monate sein. Bei der Extraktion aus Primärpublikationen wird die Anwendung eines Sättigungskriteriums von 10 (d. h. 10 konsekutive Publikationen ohne neu-identifizierte Confounder-Variable) empfohlen.

Schritt 2: Zusammenführung

Die Handlungsempfehlungen für die Zusammenführung der im 1. Schritt identifizierten Variablen sind dem Arbeitspapier GA23-02 zu entnehmen.

Schritt 3: Reduktion

Die im Arbeitspapier GA23-02 als 3. Schritt beschriebene Relevanzbewertung wird durch die folgenden Handlungsempfehlungen ersetzt: Für einen systematischen und transparenten Ansatz der Reduktion zu erhebender Confounder sollten unter Einbindung klinischer Expertinnen und Experten ein oder mehrere minimal ausreichende Adjustierungssets mittels gerichteter azyklischer Graphen erarbeitet werden. Es wird empfohlen, diese in einem Präsenz-Workshop mit einem Team aus mindestens 1 methodischen und 3 klinischen Expertinnen und Experten zu erarbeiten. Weggelassene Pfeile im Sinne eines fehlenden Einflusses auf die Therapieentscheidung oder Endpunkte müssen begründet werden. Annahmen für einen fehlenden kausalen Zusammenhang zwischen den einzelnen Confoundern müssen mindestens in tabellarischer Form transparent dokumentiert werden. Im Bericht zur Confounderidentifikation muss das finale DAG bzw. die finalen DAGs inkl. Modell-Code für das DAG aus der zur Erstellung verwendeten Software und Matrix abgebildet werden. Die Workshop-Teilnehmenden sollten namentlich benannt und potenzielle Interessenkonflikte offengelegt werden. Für die weitere Reduktion des finalen Confounder-Sets kann die Stärke des Einflusses sowohl auf die Therapieentscheidung als auch auf die Endpunkte einbezogen werden. Diese kann mithilfe einer Umfrage bei Klinikerinnen und Klinikern ermittelt werden.

Fazit

Das Verfahren der systematischen Confounderidentifikation ist ressourcenintensiv. Jedoch bestehen Möglichkeiten zur Aufwandsreduktion sowohl für die Confounderidentifikation als auch ggf. die zugehörige Erhebung der Confounder.

Anhand des RRMS-Beispiels konnte gezeigt werden, dass die Informationsbeschaffung für die Confounderidentifikation über systematische Übersichten im Vergleich zu einer Recherche nach Primärstudien grundsätzlich zeitsparender und eine Möglichkeit zur deutlichen Aufwandsreduktion sein kann. Hier ist jedoch zu beachten, dass dies nur in Situationen möglich ist, in denen über systematische Übersichten die gesamte Fragestellung abgedeckt werden kann. Die entsprechende Recherche für die Confounderidentifikation sollte bei einem

prospektiven Studiendesign zu Beginn der Datenerhebung – bzw. bei einem retrospektiven Studiendesign zum Zeitpunkt der Finalisierung des Studienprotokolls – nicht älter als 12 Monate sein.

Als geeignete Möglichkeiten zur Aufwandsreduktion bei der Erhebung von Confoundern erwiesen sich:

- die Erarbeitung eines DAG mit klinischen Expertinnen und Experten und
- die anschließende Bewertung der Relevanz der Confounder durch klinische Expertinnen und Experten.

Mithilfe von DAGs kann systematisch und transparent ermittelt werden, für welche Teilmenge der identifizierten Confounder mindestens adjustiert werden muss, um alle identifizierten Confounder adäquat in der Auswertung zu berücksichtigen. Auch wenn die für die DAG-Erarbeitung getroffenen Annahmen der Expertengruppe nicht überprüfbar sind, liefern DAGs dennoch eine transparente und nachvollziehbare Begründung für das Weglassen einzelner Confounder.

Sollte eine weitere Reduktion der Confounder-Liste notwendig sein (z. B. aufgrund einer hohen Zahl identifizierter Confounder bei nur geringer für die Analysen zur Verfügung stehender Fallzahl), wird empfohlen, diese Reduktion anhand der von klinischen Expertinnen und Experten angenommenen Stärke des Einflusses auf die Therapieentscheidung bzw. die Endpunkte vorzunehmen. Alle Confounder, die im minimal ausreichenden Adjustierungsset enthalten sind, und denen nur ein untergeordneter Einfluss beigemessen wird, können aus dem finalen Confounder-Set entfernt werden. Wird die Mehrzahl an Confoundern als wichtig eingeordnet, ist keine relevante Aufwandsreduktion möglich ohne das Risiko eines nicht abschätzbaren Verlusts an Ergebnissicherheit einzugehen.

Die beschriebenen Maßnahmen zur Aufwandsreduktion können in das im Arbeitspapier GA23-02 vorgeschlagene Verfahren zur systematischen Confounderidentifikation integriert werden. Dieses setzt sich nun aus den folgenden Schritten zusammen:

- 1) Confounderidentifikation bestehend aus einer systematischen Literaturrecherche und aus Interviews mit klinischen Expertinnen und Experten,
- 2) Zusammenführung von Confounder-Variablen, welche ein inhaltlich gleiches Konstrukt messen oder sich inhaltlich überlappen,
- 3) Reduktion des zuvor identifizierten Confounder-Sets in 2 Teilschritten:
 - a) Identifikation eines minimal ausreichenden Adjustierungssets mithilfe eines DAG und unter Einbindung klinischer Expertinnen und Experten

- b) Ausschluss von Confoundern untergeordneter Relevanz auf Basis der Einschätzung durch klinische Expertinnen und Experten (optional)

Die Durchführung dieser Schritte ist dabei vollumfänglich zu dokumentieren.

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 eine nahezu unverzerrte Effektschätzung durch die im minimal ausreichenden Adjustierungsset identifizierten und durch die Bewertung der Relevanz ausgewählten Confounder erreicht werden kann. 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 für die keine Confounderidentifikation und -kontrolle erforderlich ist. Registerbasierte RCTs verbinden dabei verschiedene Vorteile, erstens diejenigen der Randomisierung und zweitens diejenigen der Nutzung versorgungsnaher Daten in einer bestehenden Infrastruktur.

1 Hintergrund

Für die Nutzenbewertung von Arzneimitteln können auch nicht randomisierte Studiendesigns eine Rolle spielen – unter anderem, weil anwendungsbegleitende Datenerhebungen (AbDs) den gesetzlichen Vorgaben zufolge ohne Randomisierung durchzuführen sind. Ohne Randomisierung unterscheiden sich jedoch die Behandlungsgruppen oder Studienarme oft strukturell voneinander. Die Verzerrung der Studienergebnisse durch Confounder lässt sich durch statistische Verfahren reduzieren. Dafür müssen die Confounder jedoch bereits zu Studienbeginn bekannt sein und in der Studie mit hoher Qualität erhoben werden.

Im Arbeitspapier GA23-02 Systematische Confounderidentifikation in der Indikation schubförmig remittierende multiple Sklerose (RRMS) [8] wurden bereits konkrete Empfehlungen zur systematischen Confounderidentifikation erarbeitet. Trotz der im Arbeitspapier GA23-02 beschriebenen Möglichkeiten zur Aufwandseinsparung ist das Verfahren bislang ressourcenintensiv. Dies betrifft zum einen den Aufwand bei der systematischen Literaturrecherche für die Confounderidentifikation. Zum anderen kann der im Arbeitspapier GA23-02 beschriebene sensitive Ansatz zu einer hohen Anzahl an identifizierten potenziellen Confounder führen. Dies führt wiederum zu erhöhtem Aufwand für die Confounder-Erhebung in einer prospektiven Studie sowie eine höhere Anzahl an einzuschließenden Patientinnen und Patienten [9-11].

Basierend auf dem wiederkehrenden Diskurs über den Aufwand der systematischen Confounderidentifikation und den erforderlichen Ressourceneinsatz für die Erhebung soll im vorliegenden Generalauftrags-Projekt aufbauend auf dem im Arbeitspapier GA23-02 formulierten Vorschlag geprüft werden, ob weitere Möglichkeiten zur Aufwandsreduktion bestehen. Dies umfasst einerseits Möglichkeiten zur Aufwandsreduktion bei der systematischen Literaturrecherche und andererseits Möglichkeiten zur Reduktion der Zahl zu erhebender Confounder-Variablen. Die möglichen Optionen zur Aufwandsreduktion im Sinne der Machbarkeit sollen vor dem Hintergrund des potenziell damit einhergehenden Verlustes der Ergebnissicherheit diskutiert werden.

2 Fragestellung

Ziele der vorliegenden Untersuchung sind aufbauend auf dem Arbeitspapier GA23-02:

- Prüfung von Möglichkeiten zur Aufwandsreduktion bei der systematischen Literaturrecherche für die Confounderidentifikation,
- Prüfung weiterer Möglichkeiten zur Aufwandsreduktion bei der nachfolgenden Datenerhebung (z. B. in einem Register) durch die Reduktion der Zahl der für die nicht randomisierte Studie zu erhebenden Confounder,
- Ableitung von Handlungsempfehlungen zur weiteren Aufwandsreduktion bei der systematischen Identifikation und der Erhebung von Confoundern im Sinne einer Weiterentwicklung der Handlungsempfehlungen aus dem Arbeitspapier GA23-02,

jeweils unter Berücksichtigung des potenziell damit einhergehenden Verlustes der Ergebnissicherheit.

3 Projektverlauf

Im Rahmen des Generalauftrags wurde das Thema Evaluation von Maßnahmen zur Aufwandsreduktion bei der systematischen Identifikation und der Erhebung von Confoundern bearbeitet. Bearbeitungsbeginn war der 01.08.2025.

In die Bearbeitung des Projekts wurden 2 externe Sachverständige eingebunden.

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

4 Methoden

4.1 Aufwandsreduktion bei der Literaturrecherche für die systematische Confounderidentifikation

In Pufulete 2022 und GA23-02 wurden zur Identifikation von Confoundern Primärpublikationen zu RCTs und Kohortenstudien herangezogen. Ein alleiniger Rückgriff auf Sekundärliteratur (z. B. systematische Übersichten) zur Confounderidentifikation wird nicht empfohlen u. a. da die notwendigen Informationen (z. B. Adjustierungsvariablen) nicht zuverlässig berichtet werden. Es stellte sich jedoch die Frage, ob Sekundärliteratur (d. h. systematische Übersichten) zur Identifikation der Primärpublikationen und somit zur Confounderidentifikation genutzt werden kann.

Im Rahmen dieses Projektes wurde dies anhand von 2 Anwendungsbeispielen geprüft:

- Das 1. Beispiel ist die Indikation schubförmig remittierende Multiple Sklerose (RRMS). Diese wurde gewählt, da hier bereits im Projekt GA23-02 [8] eine Confounderidentifikation gemäß Pufulete 2022 [12] durchgeführt wurde.
- Das 2. Beispiel ist die Indikation rezidiertes oder refraktäres Mantelzell-Lymphom (MCL). Es handelt sich hier um ein Beispiel für eine seltene Erkrankung in der Anwendung innerhalb einer AbD. Der GBA hat 2022 eine AbD für den Wirkstoff Brexucabtagen-Autoleucel in der Behandlung von Erwachsenen mit rezidiertem oder refraktärem MCL nach 2 oder mehr systemischen Therapien, die einen Bruton-Tyrosinkinase(BTK)-Inhibitor einschließen, gefordert [13]. Der GBA nennt im Rahmen dieses Verfahrens eine konkrete Liste von Faktoren, die als relevante Confounder für die AbD berücksichtigt werden sollen [14]. Diese Liste wurde unter Berücksichtigung der durchgeführten Nutzenbewertung zu Brexucabtagen-Autoleucel in der vorliegenden Indikation, der durchgeführten Beratung zur Erstellung des Studienprotokolls und statistischen Analyseplans für die vorliegende AbD und der im Studienprotokoll des pharmazeutischen Unternehmers benannten Confounder erstellt. Obwohl diese Liste nicht auf einer systematischen Confounderidentifikation gemäß Pufulete 2022 oder GA23-02 beruht, wurde sie dennoch für das vorliegende Projekt als hinreichende Annäherung angesehen und als Referenzstandard verwendet.

Für beide Anwendungsbeispiele ist somit jeweils ein definiertes Set an potenziellen Confoundern vorhanden, welches im Weiteren jeweils als Referenzstandard zur Prüfung der Frage herangezogen werden konnte, ob der Rückgriff auf systematische Übersichten zur Identifikation von Primärpublikationen einen ausreichend sensitiven Ansatz zur Confounderidentifikation darstellt.

4.1.1 Confounderidentifikation über systematische Übersichten und die darin enthaltenen Primärstudien

Es wurden zunächst hochwertige und aktuelle systematische Übersichtsarbeiten identifiziert. Die Recherche erfolgte in MEDLINE (umfasst auch die Cochrane Database of Systematic Reviews), der International HTA Database, sowie auf den Websites des National Institute for Health and Care Excellence (NICE) und der Agency for Healthcare Research and Quality (AHRQ). Die Suchen wurden auf die letzten 5 Jahre eingeschränkt.

Selektion

Die recherchierten systematischen Übersichten wurden in einem 1. Schritt anhand ihres Titels und, sofern vorhanden, Abstracts in Bezug auf ihre potenzielle Relevanz bezüglich der Einschlusskriterien bewertet. Als potenziell relevant erachtete Übersichten wurden in einem 2. Schritt anhand ihres Volltextes auf Relevanz geprüft. Die Selektion erfolgte durch 1 Person und wurde anschließend von einer 2. Person überprüft. Diskrepanzen wurden durch Diskussion zwischen Beiden aufgelöst.

Hierbei wurden die in Tabelle 1 beschriebenen Ein- bzw. Ausschlusskriterien angewandt.

Tabelle 1: Ein- und Ausschlusskriterien für das Literaturscreening (mehrsseitige Tabelle)

	Beispiel 1 RRMS^a	Beispiel 2 MCL^b
Population ^c	Patientinnen und Patienten mit RRMS. Gegebenenfalls kann die systematische Übersicht zusätzlich zu RRMS auch andere Formen der multiplen Sklerose adressieren. In den Primärpublikationen muss der Anteil an Patientinnen und Patienten mit RRMS jedoch mindestens 90 % betragen.	Patientinnen und Patienten mit refraktärem oder rezidiertem MCL nach 2 oder mehr systemischen Therapien, die einen Bruton-Tyrosinkinase(BTK)-Inhibitor einschließen.
Intervention / Vergleichs-intervention ^d	Behandlung mit Dimethylfumarat und / oder Glatirameracetat	Behandlung mit folgenden Wirkstoffen, bzw. Kombinationen: ▪ Brexucabtagen autoleucel ▪ Bendamustin + Rituximab ▪ Bortezomib ± Rituximab ▪ Lenalidomid ± Rituximab ▪ R-CHOP (Rituximab, Cyclophosphamid, Doxorubicin, Vincristin, Prednison) ▪ VRCAp (Bortezomib, Rituximab, Cyclophosphamid, Doxorubicin, Prednison) ▪ Ibrutinib ▪ R-BAC (Rituximab + Bendamustin + Cytarabin) ▪ Temsirolimus ▪ R-FCM (Fludarabin + Cyclophosphamid + Mitoxantron + Rituximab) ▪ R-Cb (Rituximab + Chlorambucil) ▪ Venetoclax ▪ Hochdosistherapie mit allogener Stammzelltransplantation ▪ Hochdosistherapie mit autologer Stammzelltransplantation
Endpunkte	jährliche Schubrate, Behinderungsprogression, gesundheitsbezogene Lebensqualität und / oder unerwünschte Ereignisse ^e	Alle Endpunkte aus den Kategorien Mortalität, Morbidität, gesundheitsbezogenen Lebensqualität und Nebenwirkungen
Studientyp	Systematische Übersichten oder Metaanalysen. Diese sollten RCTs und / oder vergleichende Kohortenstudien (retrospektiv / prospektiv) einschließen.	Systematische Übersichten oder Metaanalysen. Diese sollten RCTs und / oder vergleichende Kohortenstudien (retrospektiv / prospektiv) einschließen. Sofern diese nicht in ausreichendem Maß vorliegen wird das Einschlusskriterium gemäß GA23-02 auf 1-armige Studien erweitert.

Tabelle 1: Ein- und Ausschlusskriterien für das Literaturscreening (mehrsseitige Tabelle)

	Beispiel 1 RRMS ^a	Beispiel 2 MCL ^b
a.	Die Ein- bzw. Ausschlusskriterien für RRMS wurden in Anlehnung an den Bericht der externen Sachverständigen zur systematischen Confounderidentifikation in der Indikation schubförmig remittierende multiple Sklerose [8] definiert.	
b.	Die Ein- bzw. Ausschlusskriterien für MCL wurden in Anlehnung an die AbD für den Wirkstoff Brexucabtagen-Autoleucel [13] definiert.	
c.	Grundsätzlich sollte die Definition der Population nah an der zu untersuchenden Fragestellung liegen [8]. Sofern Studien mit dieser Population jedoch nicht in ausreichendem Maß vorliegen, wurde das Screening mit erweiterten Einschlusskriterien erneut durchgeführt.	
d.	Dabei wurde nicht unterschieden, ob die Wirkstoffe als Intervention oder als Vergleichsintervention betrachtet werden.	
e.	Sofern Studien mit diesen Endpunkten nicht ausreichend vorliegen kann auf Studien mit Endpunkten aus den Kategorien Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Nebenwirkungen zurückgegriffen werden.	
AbD: anwendungsbegleitende Datenerhebung; MCL: Mantelzell-Lymphom; RCT: randomisierte kontrollierte Studie; RRMS: schubförmig remittierende multiple Sklerose		

Auswahl der systematischen Übersichten

Nachdem der Pool an systematischen Übersichten, der die Ein- und Ausschlusskriterien erfüllte, feststand, wurde eine weitere Auswahl nach folgenden Kriterien getroffen:

- Die systematischen Übersichten wurden hinsichtlich der Qualität der Informationsbeschaffung mittels AMSTAR-2 bewertet (Items 1, 4 und 5). Nur in dieser Hinsicht hochwertige systematische Übersichten (Items 1, 4, und 5 müssen mit ja beantwortet sein) wurden herangezogen.
- Es wurde geprüft, ob die systematischen Übersichten die gesamte Fragestellung in einem angemessenen Suchzeitraum beantworten. Falls solche übergreifenden systematischen Übersichten nicht vorlagen, konnte auf mehrere systematische Übersichten zurückgegriffen werden, die jeweils eine engere Fragestellung beantworten. Hier war aber darauf zu achten, dass durch den Einschluss von mehreren systematischen Übersichten möglichst gewährleistet war, dass die gesamte Fragestellung abgedeckt war. Die Abdeckung der Fragestellung durch die eingeschlossenen systematischen Übersichten wurde im Bericht beschrieben.
- Wenn möglich wurden systematische Übersichten herangezogen, die verschiedene Studientypen umfassen, z. B. neben RCTs auch vergleichende Kohortenstudien. Je nach Ein- und Ausschlusskriterien konnten auch andere Studientypen eingeschlossen sein. Falls eine systematische Übersicht nur einen Teil der Studientypen der Ein- und Ausschlusskriterien einschloss, konnte auf mehrere systematische Übersichten zurückgegriffen werden.

- Falls nach der Anwendung der oben genannten Punkte eine größere Anzahl an systematischen Übersichten vorlag, konnte über die Aktualität der Suche in den systematischen Übersichten weiter eingeschränkt werden.

Um eine sichere Abdeckung bez. Fragestellung, Studientyp und Aktualität zu erzielen, sollten als Richtwert insgesamt mind. 2 systematische Übersichten eingeschlossen werden. Falls eine sichere Abdeckung nicht gewährleistet wurde, wurde das weitere Vorgehen (z. B. Suche nach Primärpublikationen) beispielspezifisch diskutiert.

Nachdem der Pool an systematischen Übersichten feststand, wurden die in den Übersichten eingeschlossenen Primärpublikationen erfasst. Die identifizierten Primärpublikationen wurden zunächst anhand ihres Titels und, sofern vorhanden, Abstracts in Bezug auf ihre potenzielle Relevanz bezüglich der Einschlusskriterien bewertet. Als potenziell relevant erachtete Publikationen wurden in einem 2. Schritt anhand ihres Volltextes auf Relevanz geprüft. Die Selektion erfolgte durch 1 Person und wurde anschließend von einer 2. Person überprüft. Diskrepanzen wurden durch Diskussion zwischen beiden aufgelöst.

Extraktion von potenziellen Confounder-Variablen

Primärpublikationen, die die Ein- und Ausschlusskriterien erfüllten, wurden zur Extraktion der potenziellen Confounder-Variablen herangezogen. Neben grundlegenden Charakteristika wie Erstautor, Publikationsjahr und Studientyp wurden analog zum Vorgehen von Pufulete 2022 und den externen Sachverständigen im Arbeitspapier GA23-02 in den vergleichenden Kohortenstudien als potenzielle Confounder alle Faktoren, die zur Adjustierung in den statistischen Analysen verwendet wurden (unabhängig von der statistischen Signifikanz), extrahiert. Bei den RCTs und bei 1-armigen Studien wurden die Patientencharakteristika extrahiert. Die extrahierten Variablen wurden in ein Tabellenblatt überführt.

Je nach Anzahl der Primärpublikationen konnte nach einer bestimmten Anzahl an konsekutiven Studien ohne eine neu-hinzugekommene potenzielle Confounder-Variable die Extraktion gestoppt werden (sogenanntes Sättigungskriterium). Es wurde gemäß GA23-02 davon ausgegangen, dass die Verwendung von 10 konsekutiven Studien als Sättigungskriterium angemessen ist.

4.1.2 Effizienz der Verfahren

Gemäß GA23-02 müssen mögliche Optionen zur Aufwandsreduktion dem damit einhergehenden Informationsverlust sowie der Ergebnissicherheit gegenübergestellt werden.

4.1.2.1 Ergebnissicherheit und potenzieller Informationsverlust

Zunächst wurde ein Abgleich mit dem Referenzstandard (RRMS [8], MCL [14]) durchgeführt, um eine Abschätzung vornehmen zu können, wie vollständig die Identifikation von

Confoundern durch das hier beschriebene Verfahren ist und ob Informationsverluste bestehen.

Es wurde dargestellt, ob ein Informationsverlust besteht und unter Berücksichtigung der Erkenntnisse der in Abschnitt 4.2.2 beschriebenen Prüfungen diskutiert, ob die Bedeutung der ggf. nicht identifizierten potenziellen Confounder für die Ergebnissicherheit vertretbar ist.

Darüber hinaus wurde anhand der Ergebnisse geprüft:

- Welchen Einfluss hat die Auswahl von bestimmten Studientypen als Quelle für die Identifikation von potenziellen Confounder-Variablen, also sind andere Studientypen neben Kohortenstudien für die Identifikation wichtig?
- Welche Bedeutung hat die Aktualität der Recherche nach Primärstudien (in systematischen Übersichten oder in einer Aktualisierungsrecherche) auf die Ergebnisse zur Identifikation der potenziellen Confounder?
- Welchen Einfluss hat das Sättigungskriterium auf die Identifikation von potenziellen Confounder-Variablen?

4.1.2.2 Aufwandsreduktion

Anhand des Anwendungsbeispiels in der Indikation RRMS wurde der Aufwand für die Studienidentifikation mittels bibliografischer Recherche nach Primärstudien im Vergleich zur Studienidentifikation auf der Grundlage von systematischen Übersichten betrachtet.

Für die Abschätzung einer Größenordnung bezüglich des Aufwands wurden Trefferzahlen, Anzahl an gesichteten Volltexten und an identifizierten potenziellen Confounder betrachtet.

Der zeitliche Aufwand wurde hierbei anhand von Erfahrungswerten bzw. auf Basis von Literatur geschätzt (z. B. [15]). Darüber hinaus erfolgte eine Einordnung bzw. ein Vergleich der benötigten Ressourcen und der grundsätzlichen Machbarkeit der beiden Ansätze, potenzielle Confounder-Variablen zu identifizieren.

4.2 Aufwandsreduktion bei der Datenerhebung

4.2.1 Orientierende Informationsbeschaffung

Anhand einer orientierenden Recherche in PubMed wurde geprüft, ob in der Literatur Ansätze zur systematischen Reduktion eines Confounder-Sets basierend auf inhaltlichen Überlegungen unabhängig vom auszuwertenden Datensatz beschrieben werden. Eingeschlossen wurden Publikationen, in denen ein systematisches Vorgehen zur Reduktion der Variablen beschrieben wurde, mit dem Ziel a priori (vor Erhebung der Variablen) ein minimales Confounder-Set zu definieren. Publikationen, die allein durch eine post hoc Analyse

der in der Studie erhobenen Confounder eine Variablenselektion vornehmen, wurden ausgeschlossen.

Zusätzlich wurde eine Website-Suche nach methodischen Dokumenten zur Confounder-identifikation und -selektion der Zulassungsbehörden (European Medicines Agency [EMA] und Food and Drug Administration [FDA]) sowie Health-Technology-Assessment(HTA)-Agenturen durchgeführt.

4.2.2 Prüfung verschiedener Maßnahmen zur Aufwandsreduktion

Aufbauend auf den Untersuchungen in GA23-02 wurde geprüft, welche weiteren Möglichkeiten zur Aufwandsreduktion der Confundererhebung bestehen.

Dieser Fragestellung sollte sich zum einen empirisch über den umfangreichen Datensatz zu den identifizierten potenziellen Confounder-Variablen in der Indikation RRMS, welcher durch externe Sachverständige in GA23-02 [8] erstellt wurde, genähert werden. So wurde anhand des Datensatzes geprüft, ob sich Kriterien ableiten lassen, auf deren Grundlage Confounder-Variablen identifiziert werden können, auf deren Erhebung verzichtet werden kann. Dazu zählten unter anderem die Identifikation von Schwellenwerten hinsichtlich der Nennung in der Literatur oder die Berücksichtigung der Art der Quelle, aus der ein potenzieller Confounder identifiziert wurde.

Zum anderen wurden die in der Anhörung zum vorläufigen Rapid Report A25-13 [9] vorgebrachten Anregungen dahin gehend geprüft, ob sie weitere relevante Maßnahmen zur Aufwandsreduktion darstellen. Diese Anregungen umfassten:

- Berücksichtigung der Qualität der Quelle,
- Berücksichtigung von Korrelationen,
- Fokussierung auf Faktoren, die die Therapieentscheidung beeinflussen,
- Ausschluss instrumenteller Variablen und
- Anwendung gerichteter azyklischer Graphen (DAGs).

Es wurde untersucht, ob und wie diese Anregungen für die Selektion von Confoundern berücksichtigt werden können. Hierzu wurde zunächst eine Definition der einzelnen Anregungen festgehalten. Anschließend wurde ein schrittweises Vorgehen skizziert, welches notwendig wäre, um die jeweilige Anregung umzusetzen. Dabei wurde auch auf die potenziellen Limitationen der einzelnen Anregungen in Hinblick auf Informationsverlust beziehungsweise Ergebnissicherheit eingegangen. Im nachfolgenden Schritt wurden die jeweiligen Anregungen dahin gehend bewertet, ob diese als geeignete oder nicht geeignete Maßnahme zur Selektion eingeschätzt werden.

Die über die oben genannten Schritte als geeignet identifizierten Maßnahmen wurden auf die in GA23-02 identifizierten 28 Confounder angewendet, um zu überprüfen inwieweit eine weitere Reduktion der Zahl zu erhebenden Confounder erreicht werden kann. Anhand dieses Beispiels wurde zum einen diskutiert, welche Auswirkungen des Vorgehens auf die Ergebnissicherheit zu erwarten sind. Zum anderen wurde diskutiert werden, inwiefern ein höherer Aufwandseinsatz zur Reduktion der Anzahl zu erhebender Variablen vor Beginn der Datenerhebung im Abgleich mit den Aufwandseinsparungen bei der Datenerhebung und -auswertung gerechtfertigt ist.

4.3 Ableitung von Anforderungen an eine systematische Confounderidentifikation (Handlungsempfehlungen)

Anhand der gesichteten Literatur und den Ergebnissen der Untersuchungen anhand des MS-Datensatzes wurden die konkreten Handlungsempfehlungen zur systematischen Identifikation und der Erhebung von Confoundern für nicht randomisierte vergleichende Studien aus GA23-02 im Sinne der Aufwandsreduktion weiterentwickelt. Hierbei wurde im Speziellen geprüft, inwieweit sich für Fragestellungen mit erwartbar weniger umfassender Studienlage (wie beispielsweise Fragestellungen der AbD) besondere Anforderungen ergeben.

5 Ergebnisse

5.1 Aufwandsreduktion bei der Literaturrecherche

5.1.1 Confounderidentifikation über systematische Übersichten und die darin enthaltenen Primärstudien

Im Folgenden wird das Ergebnis der Confounderidentifikation über systematische Übersichten getrennt für die Beispiele RRMS und MCL beschrieben.

5.1.1.1 Beispiel RRMS

5.1.1.1.1 Ergebnisse der Informationsbeschaffung

Suche und Identifikation von qualitativ hochwertigen und aktuellen systematischen Übersichten

Abbildung 1 zeigt das Ergebnis der systematischen Literaturrecherche und der Studieneinklassifizierung gemäß den Kriterien für den Studieneinschluss. Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich im Anhang A.1.1.1. Die letzte Suche fand am 04.07.2024 statt.

Die Referenzen der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Anhang C.1.2.

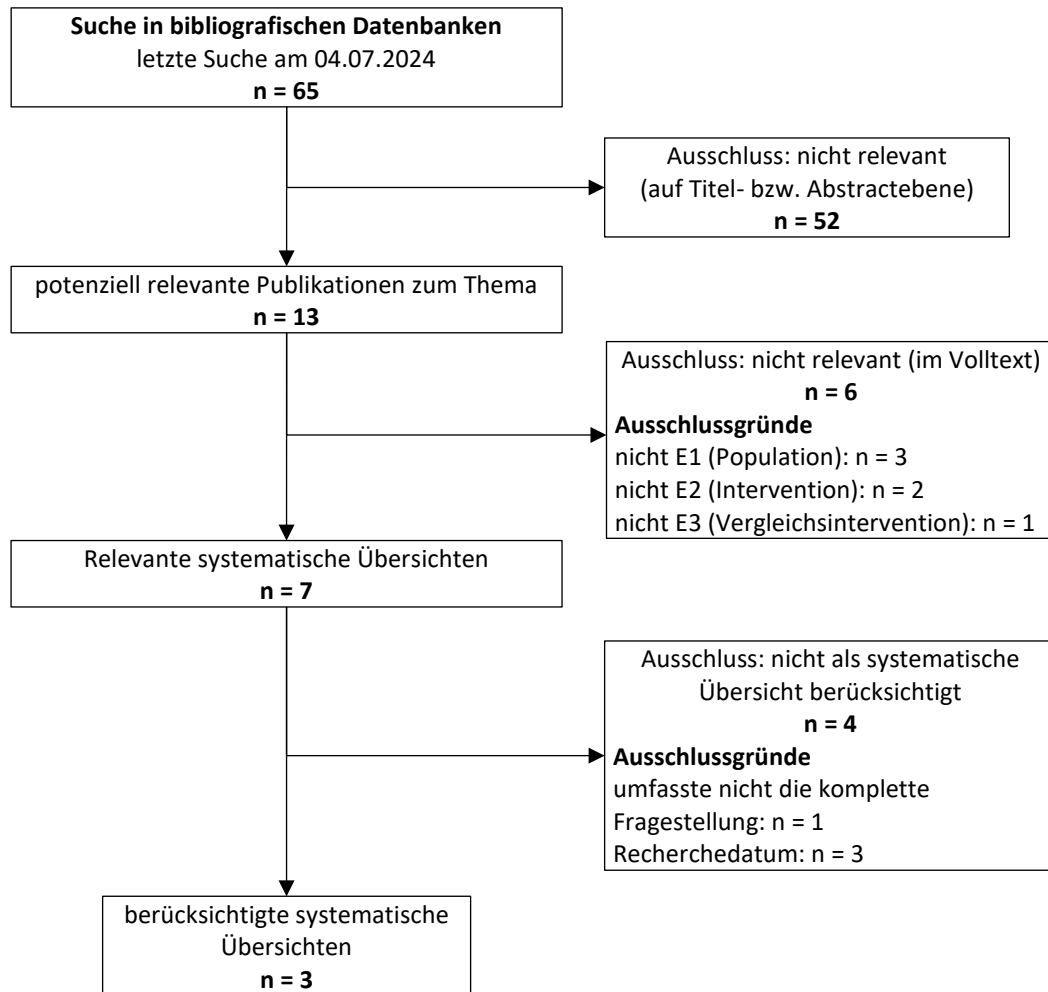


Abbildung 1: Ergebnis der Informationsbeschaffung und der Studienselektion nach systematischen Übersichten (Beispiel RRMS)

Berücksichtigte systematische Übersichten

Von den 7 nach den Einschlusskriterien als relevant eingeordneten systematischen Übersichten (siehe Anhang C.1.1) wurden 3 berücksichtigt, die zum einen inhaltlich den größtmöglichen Teil der Fragestellung abdecken und zum anderen eine ausreichende Qualität aufweisen (mit AMSTAR-2 als hochwertig eingestuft), beide Studientypen (RCTs und vergleichende Kohortenstudien) abdecken und ein möglichst aktuelles Recherchedatum aufweisen. Eine systematische Übersicht wurde nicht berücksichtigt, da sie die Fragestellung hinsichtlich der Indikation nur eingeschränkt umfasst [16]. Weitere 3 systematische Übersichten wurden aufgrund des aktuelleren Recherchedatums [17-19] der bereits eingeschlossenen systematischen Übersichten nicht weiter betrachtet, hätten darüber hinaus jedoch auch keinen zusätzlichen Vorteil erbracht (z. B. einen zusätzlichen Teil der Fragestellung abgedeckt).

Eine Übersicht der berücksichtigten systematischen Übersichten ist in Tabelle 2 dargestellt. Die Bewertung der Qualität der Informationsbeschaffung dieser systematischen Übersichten findet sich in Anhang B.1.

Tabelle 2: Berücksichtigte systematische Übersichten (Beispiel RRMS)

Systematische Übersicht [Zitat]	Recherche-datum	Zeitlicher Unterschied zur Erstellung Referenzstandards	RCT	vergleichende Kohortenstudien	Betrachtete Interventionen
Gonzalez-Lorenzo 2024 [20]	08/2022	< 1 Jahr	ja	nein	▪ Glatirameracetat ▪ Dimethylfumarat
Tramacere 2023 [21]	03/2022	1 Jahr, 7 Monate	ja	nein	▪ Glatirameracetat ▪ Dimethylfumarat
Prosperini 2023 [22]	12/2022	< 1 Jahr	nein	ja	▪ Dimethylfumarat
RCT: randomisierte kontrollierte Studie; RRMS: schubförmig remittierende multiple Sklerose					

Die 3 ausgewählten systematischen Übersichten decken nicht die gesamte Fragestellung ab. Es fehlen vergleichende Kohortenstudien, die Studien mit Glatirameracetat einschließen. Aus diesem Grund ist entsprechend der Ein- und Ausschlusskriterien für den nicht abgedeckten Teil der Fragestellung eine ergänzende Recherche nach Primärpublikationen notwendig. Hierfür wurde auf den im Arbeitspapier GA23-02 [8] veröffentlichten Datensatz der externen Sachverständigen zurückgegriffen. Die Referenzen der in diesem Datensatz enthaltenen Studien finden sich in Anhang C.1.5. Im Folgenden wird bei der Ergebnisdarstellung jeweils differenziert zwischen den ausschließlich über systematischen Übersichten identifizierten Primärpublikationen und dem Ergebnis inklusive ergänzender Recherche nach Primärpublikationen.

Identifizierte Primärpublikationen

Abbildung 2 zeigt das Ergebnis der Sichtung der eingeschlossenen Publikationen in den 3 ausgewählten systematischen Übersichten gemäß den Kriterien für den Studieneinschluss. In der Abbildung sind ebenfalls die Primärpublikationen aus dem Datensatz der externen Sachverständigen aus dem Arbeitspapier GA23-02 aufgeführt, d. h. vergleichende Kohortenstudien, die Glatirameracetat als Intervention betrachten. Die Referenzen der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Anhang C.1.4

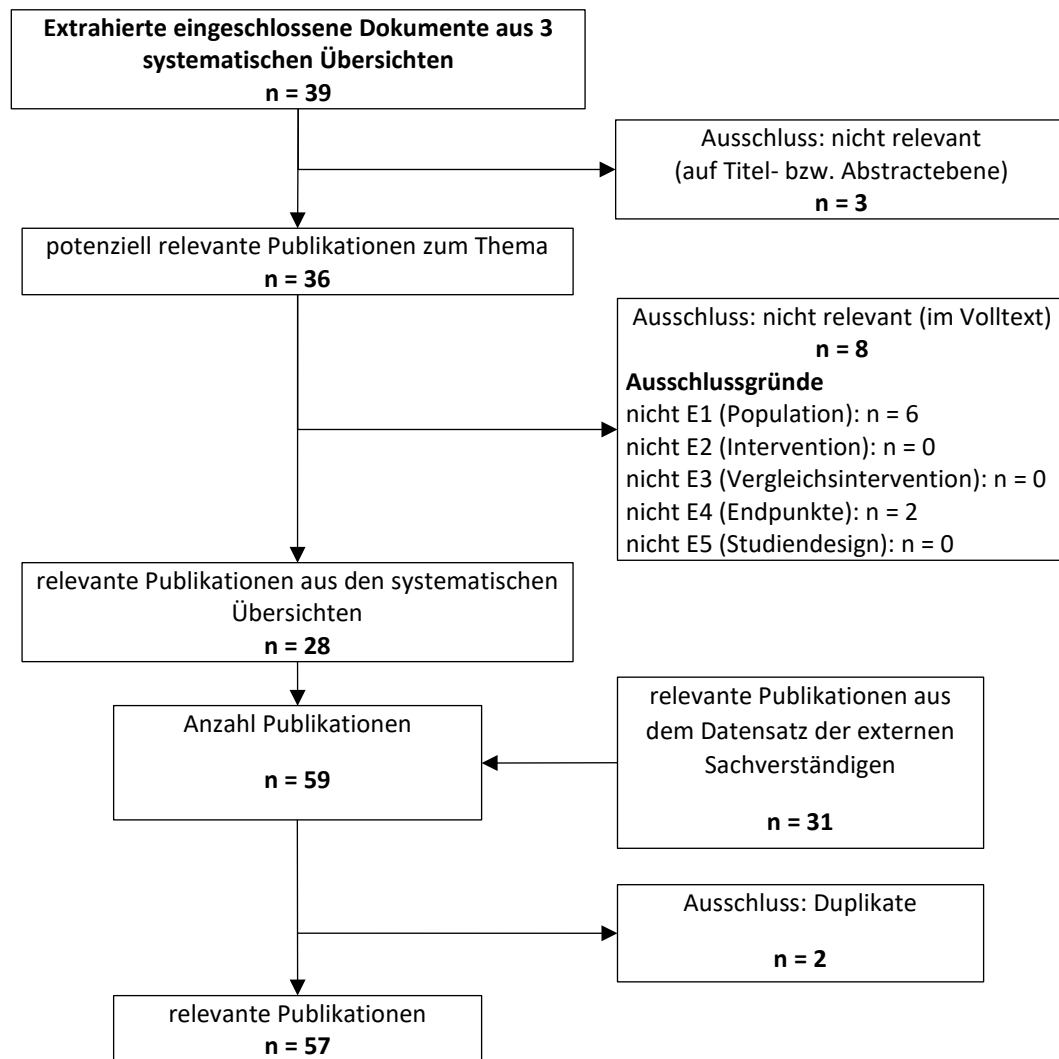


Abbildung 2: Ergebnis der Sichtung der eingeschlossenen Publikationen (Beispiel RRMS)

Für die relevanten 57 Publikationen wurde kein Sättigungskriterium angewendet.

Studienpool

Insgesamt wurden aus den 3 systematischen Übersichten 28 Primärpublikationen eingeschlossen, davon 17 RCTs und 11 Kohortenstudien. Die Liste der eingeschlossenen Primärpublikationen aus den systematischen Übersichten findet sich in Anhang C.1.3.

Aus dem veröffentlichten Datensatz der externen Sachverständigen wurden nach Dublettenabgleich 29 Primärpublikationen ausschließlich zu Kohortenstudien und zu Glatirameracetat eingeschlossen. Die Referenzen der in diesem Datensatz enthaltenen Referenzen finden sich in Anhang C.1.5.

Für die Extraktion der Confounder-Variablen liegen somit als Grundlage 57 Primärpublikationen vor.

5.1.1.1.2 Ergebnisse der Auswertung

Aus den 28 in den systematischen Übersichten enthaltenen Primärpublikationen wurden insgesamt 49 Confounder-Variablen extrahiert. Aus dem veröffentlichten Datensatz der externen Sachverständigen wurden insgesamt 42 Confounder-Variablen übernommen.

Abgleich mit Referenzstandard

Die identifizierten Confounder-Variablen aus den Primärpublikationen aus den systematischen Übersichten und dem veröffentlichten Datensatz der externen Sachverständigen wurden mit einem Referenzstandard abgeglichen, der für die gleiche Fragestellung auf Grundlage einer Confounderidentifikation ausschließlich über Primärpublikationen (gemäß Pufulete 2022 [12]) erstellt wurde [8]. Die Ergebnisse finden sich in Tabelle 3.

Hier ist zu beachten, dass in den Referenzstandard potenzielle Confounder sowohl durch die Literaturrecherche als durch Interviews mit Expertinnen und Experten eingeflossen sind. 2 der 28 Confounder wurden ausschließlich durch die Interviews identifiziert. Da im vorliegenden Arbeitspapier in Abschnitt 5.1 nur die Aufwandsreduktion durch die Literaturrecherche betrachtet wird, werden diese beiden nur im Interview identifizierten Confounder in der Auswertung nicht berücksichtigt. Die Auswertung bezieht sich somit insgesamt nur auf 26 potenzielle Confounder des Referenzsets.

Tabelle 3: Abgleich mit Referenzstandard (Beispiel RRMS) (mehrseitige Tabelle)

Potenzieller Confounder aus GA23-02 (Referenzstandard)	Identifiziert durch Primärpublikationen aus SÜ	Identifiziert durch Primärpublikationen aus externem Datensatz ^a	Identifiziert durch Interview im Rahmen von GA23-02 ^a
Geschlecht	ja	–	–
Alter	ja	–	–
EDSS Score	ja	–	–
EDSS Veränderung im Vorjahr	unklar ^b	ja	unklar ^b
Multiple Sclerosis Functional Composite (MSFC) Score	ja	–	–
Fatigue	nein	ja	nein
Komorbiditäten	ja	–	–
absolute Lymphozytenzahl; Lymphopenie	nein	nein ^c	ja
Zentrum für Multiple Sklerose	ja	–	–
BMI	ja	–	–
Tabakkonsum	ja	–	–
infratentorielle Läsionen	nein	ja	ja
Anzahl der Kontrastmittel- anreichernden Läsionen	ja	–	–
Volumen T1-hypointenser Läsionen	ja	–	–
Anzahl der T2-Läsionen	ja	–	–
Volumen der T2-Läsionen	ja	–	–
Verträglichkeit der Vortherapie und Erwartungen an die Behandlung	nein	ja	ja
vorherige Behandlung mit einer anderen krankheitsmodifizierenden Therapie	ja	–	–
Patientenwunsch ^d	nein	nein	ja
gesundheitsbezogene Lebensqualität	ja	–	–
Anzahl der vorherigen Schübe	ja	–	–
Schwere der Schübe	ja	–	–
Zeit seit der Diagnose der Multiplen Sklerose	ja	–	–
Cerebrospinalflüssigkeit (CSF) Status	nein	ja	ja
Art der beruflichen Tätigkeit ^d	nein	nein	ja
Bildungsgrad	nein	nein ^c	nein
Abstammung	ja	–	–
kürzliche Schwangerschaft (in den letzten 3 Monaten)	nein	ja	nein

Tabelle 3: Abgleich mit Referenzstandard (Beispiel RRMS) (mehrseitige Tabelle)

Potenzieller Confounder aus GA23-02 (Referenzstandard)	Identifiziert durch Primärpublikationen aus SÜ	Identifiziert durch Primärpublikationen aus externem Datensatz ^a	Identifiziert durch Interview im Rahmen von GA23-02 ^a
a. Hier wurden jeweils nur die Variablen aufgeführt, die nicht eindeutig durch die Primärpublikationen aus SÜs identifiziert wurden. b. Es ist unklar, ob der potenzielle Confounder identifiziert wurde. Es wurden zwar ähnliche Confounder-Variablen genannt, jedoch nicht der potenzielle Confounder direkt (z. B. wurde zwar der EDSS Score genannt, jedoch nicht die Veränderung im Vorjahr). c. bezieht sich nur auf den hier betrachteten Datensatz, d. h. vergleichende Kohortenstudien, die Glatirameracetat als Intervention betrachten d. Dieser potenzielle Confounder wurde im Referenzset nur durch das Interview identifiziert. Da im vorliegenden Arbeitspapier in Abschnitt 5.1 nur die Aufwandsreduktion durch die Literaturrecherche betrachtet wird, wird dieser potenzielle Confounder in den folgenden Tabellen nicht aufgeführt. BMI: Body Mass Index; EDSS: Expanded Disability Status Scale; MS: multiple Sklerose; RRMS: schubförmig remittierende multiple Sklerose; SÜ: systematische Übersicht			

Die Anzahl der identifizierten potenziellen Confounder in Abhängigkeit des Datensatzes ist in Tabelle 4 dargestellt.

Tabelle 4: Anzahl der identifizierten potenziellen Confounder aus dem Referenzset (N = 26) in Abhängigkeit des Datensatzes (Beispiel RRMS)

Potenzieller Confounder identifiziert?	Identifiziert durch Primärpublikationen aus SÜ	Identifiziert durch Primärpublikationen aus SÜ + externem Datensatz	Identifiziert durch Primärpublikationen aus SÜ + Interview	Identifiziert durch Primärpublikationen aus SÜ + externem Datensatz + Interview
ja	18	24	22	25
unklar ^a	1	0	1	0
nein	7	2	3	1
a. Es ist unklar, ob der potenzielle Confounder identifiziert wurde. Es wurden zwar ähnliche Confounder-Variablen genannt, jedoch nicht der potenzielle Confounder direkt (z. B. wurde zwar der EDSS score genannt, jedoch nicht die Veränderung im Vorjahr). EDSS: Expanded Disability Status Scale; RRMS: schubförmig remittierende multiple Sklerose; SÜ: systematische Übersicht				

Durch die in den 3 systematischen Übersichten enthaltenen Primärpublikationen werden etwa 2 Drittel (d. h. 18) der 26 potenziellen Confounder identifiziert. Durch die Berücksichtigung der Kohortenstudien zu Glatirameracetat aus dem Datensatz der externen Sachverständigen [8] wurden 6 weitere, d. h. insgesamt 24 potenzielle Confounder identifiziert. Durch die zusätzliche Berücksichtigung der im Interview genannten Variablen wurde wiederum 1 weiterer potenzieller Confounder, d. h. insgesamt 25 potenzielle Confounder, identifiziert. Zusammenfassend bedeutet dies, dass selbst mit der ergänzenden

Recherche nach Primärpublikationen 1 potenzieller Confounder (Bildungsgrad) nicht identifiziert werden konnte.

Bedeutung des Informationsverlustes für die Ergebnissicherheit

Welche Bedeutung die (Nicht-)Berücksichtigung des Confounders Bildungsgrad für die Ergebnissicherheit hat, lässt sich nicht sicher abschätzen. Als grobe Annäherung hierfür kann das Ergebnis des in Abschnitt 5.2.1.8 beschriebenen Umfrage herangezogen werden. In dieser wurde die Bedeutung des Confounders Bildungsgrad von den befragten klinischen Expertinnen und Experten für die Therapieentscheidung im Median mit 1,0 (sehr gering) und für den Endpunkt mit 2,0 (gering) bewertet. Auch wenn dies nur auf der Einschätzung auf einer kleinen Stichprobe an klinischen Expertinnen und Experten beruht, wird dennoch der Informationsverlust als gering eingeschätzt.

Studientyp

Anhand der Ergebnisse wurde weiterhin geprüft, welchen Einfluss die Auswahl von bestimmten Studientypen als Quelle auf die Identifikation von potenziellen Confounder-Variablen hat, also ob andere Studientypen neben Kohortenstudien für die Identifikation wichtig sind. Dazu wurden die Primärpublikationen unterteilt in Kohortenstudien (Extraktionsquelle: berichtete Adjustierungsfaktoren) und in RCT (Extraktionsquelle: Patientencharakteristika). Tabelle 5 zeigt die Ergebnisse der Auswertung.

Tabelle 5: Anzahl der identifizierten potenziellen Confounder aus dem Referenzset (N = 26) in Abhängigkeit des Studientyps bzw. der Extraktionsquelle^a (Beispiel RRMS)

Potenzieller Confounder identifiziert?	aus Kohortenstudien (n = 40)	aus RCT (n = 17)
ja	21	15
unklar	0	0
nein	5	11
a. In den Kohortenstudien wurden die Confounder-Variablen aus den berichteten Adjustierungsfaktoren extrahiert. In den RCTs wurden die Patientencharakteristika extrahiert. RCT: randomisierte kontrollierte Studie; RRMS: schubförmig remittierende multiple Sklerose		

Durch das alleinige Betrachten der berichteten Adjustierungsfaktoren in Kohortenstudien wurden 21 der 26 potenziellen Confounder identifiziert. Durch die zusätzliche Berücksichtigung der im Interview genannten Variablen wurde 1 weiterer potenzieller Confounder identifiziert, d. h. insgesamt 22 potenzielle Confounder identifiziert.

Durch das alleinige Betrachten der in RCT genannten Patientencharakteristika wurden 15 der 26 potenziellen Confounder identifiziert. Durch die zusätzliche Berücksichtigung der im Interview genannten Variablen wurden 5 weitere potenzielle Confounder identifiziert, d. h. insgesamt 20 potenzielle Confounder identifiziert.

Zusammenfassend zeigt sich beim Beispiel RRMS, dass durch die Beschränkung auf Kohortenstudien ein Informationsverlust entsteht, welcher nicht über die zusätzliche Berücksichtigung der Interviews ausgeglichen wird.

Aktualität der Recherche

Die Bedeutung der Aktualität der Recherche nach Primärstudien (in systematischen Übersichten oder in einer Aktualisierungsrecherche) auf die Ergebnisse zur Identifikation der potenziellen Confounder wurde beim Beispiel RRMS anhand von 2 unterschiedlichen (teilweise überlappenden) Datensätzen untersucht. Zum einen wurde anhand der Primärpublikationen aus den systematischen Übersichten und der ergänzenden Recherche die Bedeutung der Aktualität untersucht. Zum anderen wurde anhand des vollständigen Datensatzes von Primärpublikationen, der Grundlage für die Confounderidentifikation der externen Sachverständigen in GA23-02 war, geprüft, in welchem Publikationsjahr potenzielle Confounder-Variablen identifiziert wurden.

Bedeutung der Aktualität anhand der Primärpublikationen aus den systematischen Übersichten und dem zusätzlichen Datensatz der externen Sachverständigen

Es wurde geprüft, welchen Einfluss die Aktualität der Primärstudien auf die Anzahl der identifizierten potenziellen Confounder hat. Dazu wurden von den eingeschlossenen Primärpublikationen sukzessive die aktuellsten Publikationen jahrweise aus der Auswertung ausgeschlossen. Tabelle 6 zeigt die Ergebnisse der Auswertung.

Tabelle 6: Anzahl der identifizierten potenziellen Confounder aus dem Referenzset (N = 26) in Abhängigkeit der Aktualität (identifiziert durch Primärpublikationen aus systematischen Übersichten + externem Datensatz)

Potenzieller Confounder identifiziert?	alle Publikationen bis einschließlich 2023 ^a (n = 57)	Publikationen bis einschließlich 2022 (n = 53)	Publikationen bis einschließlich 2021 (n = 49)	Publikationen bis einschließlich 2020 (n = 39)
ja	24	23	23	21
unklar	0	1	1	0
nein	2	2	2	5
a. Aktuellstes Recherchedatum in der systematischen Übersicht ist jeweils das Jahr 2022. Durch den Datensatz der Externen sind dennoch Publikationen von 2023 enthalten.				

Die Recherche nach systematischen Übersichten erfolgte 07/2024; die Recherche nach Primärstudien der Externen erfolgte 11/2023. Durch das Ausschließen der Publikationen aus dem Jahr 2023 und 2022 wird jeweils 1 potenzieller Confounder weniger identifiziert. Der Ausschluss der Publikationen aus dem Jahr 2021 führt zu 3 potenziellen Confoundern weniger.

Bedeutung der Aktualität anhand des vollständigen Datensatzes der externen Sachverständigen

Anhand des vollständigen Datensatzes der externen Sachverständigen wurde ebenfalls die Bedeutung der Aktualität überprüft. Durch die externen Sachverständigen wurden in der Literaturrecherche 125 potenzielle Confounder-Variablen aus 80 Publikationen extrahiert. Die Identifikation einer Confounder-Variablen kann somit einem Erscheinungsjahr der Publikation zugeordnet werden.

Im Arbeitspapier zu GA23-02 [8] wurden duplizierte bzw. stark überlappende Confounder-Variablen zu potenziellen Confoundern zusammengefasst, sodass sich aus den 125 Variablen insgesamt 26 potenzielle Confounder ergaben. Um zu untersuchen, wann, d. h. in welchem Publikationsjahr, die jeweiligen potenziellen Confounder identifiziert wurden, wurde zur Vereinfachung die erste Nennung einer Confounder-Variablen des zugeordneten potenziellen Confounders jeweils als identifiziert gewertet.

Tabelle 7 zeigt die identifizierten potenziellen Confounder und deren Auswertung hinsichtlich erstmaliger Nennung.

Tabelle 7: Identifizierte potenzielle Confounder ausgewertet nach erstmaliger Nennung

Jahr	Identifizierte potenzielle Confounder ^a	Anzahl der identifizierten potenziellen Confounder
1998	- Alter - EDSS Score - Zeit seit der Diagnose der Multiplen - Anzahl der vorherigen Schübe	4
2001	- Geschlecht - Zentrum für Multiple Sklerose - Anzahl der Kontrastmittel-anreichernden Läsionen - Volumen der T2-Läsionen - MS Functional Composite (MSFC) Score - EDSS Veränderung im Vorjahr	6
2003	- Anzahl der T2-Läsionen - Volumen T1-hypointenser Läsionen	2
2005	vorherige Behandlung mit einer anderen krankheitsmodifizierenden Therapie	1
2006	Schwere der Schübe	1
2009	Abstammung	1
2012	BMI	1
2014	- Komorbiditäten - Fatigue	2
2015	- gesundheitsbezogene Lebensqualität - Cerebrospinalflüssigkeit (CSF) Status - Verträglichkeit der Vorthherapie und Erwartungen an die Behandlung	3
2019	Tabakkonsum	1
2020	absolute Lymphozytenzahl; Lymphopenie	1
2021	- infratentorielle Läsionen - Bildungsgrad - kürzliche Schwangerschaft (in den letzten 3 Monaten)	3
a. Die potenziellen Confounder sind aus insgesamt 125 Confounder-Variablen entstanden, indem duplizierte bzw. stark überlappende Confounder-Variablen zu 26 potenziellen Confoundern zusammengeführt wurden. Es zählt die erste Nennung einer Confounder-Variablen des zugeordneten potenziellen Confounders als identifiziert. BMI: Body-Mass-Index; EDSS: Expanded Disability Status Scale; MS: multiple Sklerose		

Die Recherche nach Primärstudien der Externen erfolgte 11/2023. Bis zum Jahr 2021 wurden alle 26 potenziellen Confounder aus 63 Publikationen extrahiert. In den Jahren 2022 und 2023 wurde kein weiterer potenzieller Confounder identifiziert, obwohl weitere 17 Publikationen berücksichtigt wurden. Insgesamt ist über den zeitlichen Verlauf keine gleichmäßige Akkumulation bzw. Sättigung der identifizierten potenziellen Confounder zu erkennen.

Zeitlicher Aufwand

Zur Abschätzung des zeitlichen Aufwands wurde die benötigte Zeit für die Literaturrecherche, das Screening und die darauffolgenden Extraktionsschritte anhand der Trefferzahlen und der in den darauffolgenden Screeningschritten jeweils eingeschlossenen Publikationen anhand von Erfahrungs- und Literaturwerten [15] geschätzt und addiert. Hierzu wurde beispielsweise im Titel- / Abstract-Screening angenommen, dass innerhalb einer Stunde 50 Referenzen gesichtet werden können. Mithilfe dieser Annahmen wurde die hier dargestellte Confounderidentifikation über systematische Übersichten dem Verfahren der externen Sachverständigen – also einer Recherche nach Primärpublikationen (siehe GA23-02 [8]) – gegenübergestellt.

Wie beschrieben, wurde für die Abdeckung der Fragestellung in Bezug auf Kohortenstudien, die Glatirameracetat untersuchen, auf den Datensatz der externen Sachverständigen zurückgegriffen. Die Suche der externen Sachverständigen bezog sich jedoch auf eine andere (weitergefasste) Fragestellung. Um abzuschätzen, wie der Aufwand für die zusätzliche Suche nach Kohortenstudien, die Glatirameracetat untersuchen, war, wurde eine Informationsbeschaffung simuliert. Hierfür wurde näherungsweise eine Suche in MEDLINE durchgeführt. Andere Datenbanken, insbesondere Embase wurden nicht durchsucht, da frühere Untersuchungen keinen relevanten Mehrwert für Embase bei der Suche nach Arzneimitteln zeigten [23]. Ebenfalls wurde auf eine Suche in einer weiteren Datenbank (z. B. CENTRAL) verzichtet, da MEDLINE den Großteil der Treffer abbildet und es hier nur um eine Annäherung der Trefferzahlen geht. Die resultierende Trefferzahl aus der MEDLINE-Suche wurde als Schätzung für den Aufwand herangezogen, es wurde jedoch kein Screening durchgeführt. Die Suchstrategie findet sich in Anhang A.1.1.2.

Die Ergebnisse sind in Tabelle 8 dargestellt.

Tabelle 8: Vergleich der Aufwände bei verschiedenen Vorgehensweisen der Informationsbeschaffung und -auswertung (Beispiel RRMS)

Vorgehen	Über SÜs	Simulierte ergänzende Recherche nach Primärstudien	Primärstudien-suche der externen Sachverständigen
Trefferzahlen			
Suche in	MEDLINE, HTA Database, Handsuche (SÜs)	MEDLINE	MEDLINE, Embase
SÜs Titel / Abstract	65	-	-
SÜs Volltext	13	-	-
SÜ eingeschlossen / SÜ berücksichtigt	7 / 3	-	-
Primärstudien Titel / Abstract	39	473 ^a	802 ^b
Primärstudien Volltext	36	116 ^c	196
Primärstudien eingeschlossen	28	47 ^c	80
Gesamtaufwand ^d			
Suche ^e	15 Std.	15 Std. ^f	20 Std.
Titel- / Abstract-Screening (Anzahl; Aufwand, wenn 50 pro 1 Std. gescreent werden können)	104 Referenzen; 2 Std.	473 Referenzen; 9 Std.	802 Referenzen; 16 Std.
Volltext-Screening (Anzahl; 6 pro 1 Std.)	49 Volltexte; 8 Std.	116 ^c Volltexte; 19 Std.	196 Volltexte; 33 Std.
Extraktion der Confounder-Variablen (Anzahl; Aufwand, wenn 2 Publikationen in 1 Std. bearbeitet werden können)	28 Publikationen; 14 Std.	47 ^c Publikationen; 24 Std.	80 Publikationen; 40 Std.
Gesamt	39 Stunden	67 Stunden	109 Stunden
	107 Stunden		

a. Die Suche wurde durchgeführt und die resultierende Trefferzahl wurde als Schätzung für den Aufwand herangezogen, es wurde jedoch kein Screening durchgeführt. Die Suchstrategie findet sich in Anhang A.1.1.2.

b. 1202 Treffer; 802 tatsächlich gescreent

c. Die Zahl ist geschätzt, da dieses Screening nicht durchgeführt wurde. Hier wurde lediglich das Verhältnis der Treffer der Suche zu den gescreenten und extrahierten Publikationen der externen Sachverständigen auf das Vorgehen über SÜs übertragen.

d. Aufwand wurde anhand von Erfahrungs- und Literaturwerten [15] geschätzt, um die Vorgehensweisen miteinander zu vergleichen. Es handelt sich nicht um tatsächliche Aufwände. Qualitätssicherungsschritte sind jeweils enthalten (z. B. 4-Augen-Prinzip)

e. grobe Schätzung; Aufwand der Suche ist u. a. abhängig von der Trefferzahl und der Anzahl der durchsuchten Datenbanken

f. Für die Simulation wurde nur eine Datenbank durchsucht, weshalb 15 Std. geschätzt wurden. Bei einer auf Vollständigkeit angelegten Suche würde sich der Aufwand durch eine 2. Datenbank voraussichtlich auf 20 Std. erhöhen.

HTA: Health Technology Assessment; RRMS: schubförmig remittierende multiple Sklerose; Std.: Stunde; SÜ: systematische Übersicht

Es zeigt sich, dass durch eine ergänzende Recherche nach Primärpublikationen keine Aufwandsreduktion möglich ist (107 vs. 109 Stunden). Ist ein Rückgriff auf systematische Übersichten zur Extraktion von Primärpublikationen möglich, kann der Aufwand wahrscheinlich erheblich reduziert werden (39 vs. 109 Stunden).

5.1.1.2 Beispiel MCL

5.1.1.2.1 Ergebnisse der Informationsbeschaffung

Suche und Identifikation von qualitativ hochwertigen und aktuellen systematischen Übersichten

Abbildung 3 zeigt das Ergebnis der systematischen Literaturrecherche und der Studienselektion gemäß den Kriterien für den Studieneinschluss. Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Abschnitt A.1.2.1. Die letzte Suche fand am 21.08.2024 statt.

Die Referenzen der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Abschnitt C.2.2.

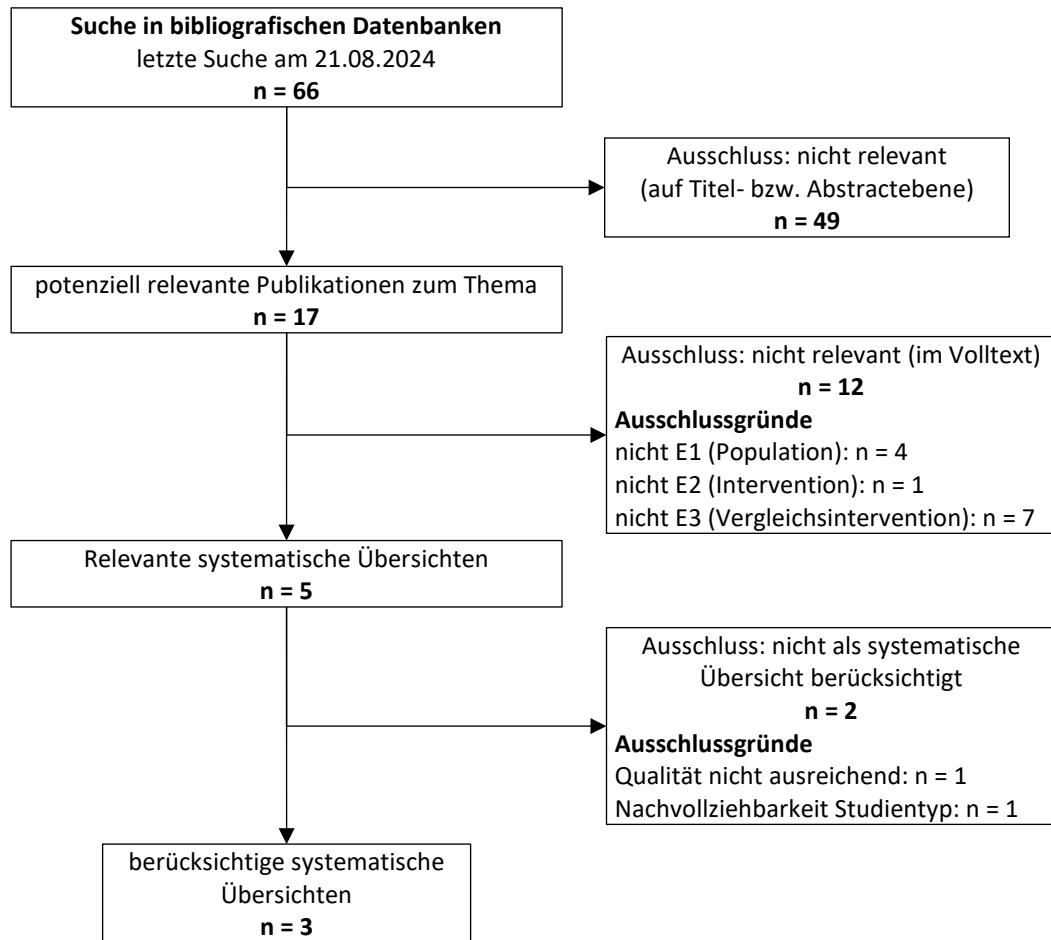


Abbildung 3: Ergebnis der Informationsbeschaffung und der Studienselektion nach systematischen Übersichten (Beispiel MCL)

Berücksichtigte systematische Übersichten

Grundsätzlich sollten die Einschlusskriterien für die Population nah an der zu untersuchenden Fragestellung liegen [8]. Die in diesem Beispiel in Anlehnung an die AbD für den Wirkstoff Brexucabtagen-Autoleucel definierte Population [13] umfasst Patientinnen und Patienten mit refraktärem oder rezidiertem MCL nach 2 oder mehr systemischen Therapien, die einen Bruton-Tyrosinkinase(BTK)-Inhibitor einschließen. Da Studien mit dieser Population jedoch nicht in ausreichendem Maß vorliegen, wurde das Screening mit erweiterten Einschlusskriterien erneut durchgeführt. Das Einschlusskriterium für die Population war hierbei wie folgt formuliert: Patientinnen und Patienten mit refraktärem oder rezidiertem MCL.

Bezüglich des Studientyps wurde nach systematischen Übersichten oder Metaanalysen gesucht. Diese sollten RCTs und / oder vergleichende Kohortenstudien (retrospektiv / prospektiv) einschließen. Insgesamt wurden jedoch über die 3 systematischen Übersichten nur 1 RCT und 1 Kohortenstudie identifiziert. Die Anzahl an Primärpublikationen wurde als

nicht ausreichend erachtet, weshalb das Einschlusskriterium gemäß GA23-02 auf 1-armige Studien erweitert wurde.

Von den 5 als relevant eingeordneten systematischen Übersichten wurden 3 berücksichtigt. Roufarshbaf 2022 [24] wurde nicht berücksichtigt, da aus der Publikation nicht klar hervorging, welcher Studientyp eingeschlossen wurde. Monga [25] wurde ebenfalls nicht berücksichtigt, da diese systematische Übersicht gemäß AMSTAR-2 nicht als hochwertig bewertet wurde.

Eine Übersicht der berücksichtigten systematischen Übersichten ist in Tabelle 9 dargestellt. Die Bewertung der Qualität der Informationsbeschaffung dieser systematischen Übersichten findet sich in Anhang B.2.

Tabelle 9: Berücksichtigte systematische Übersichten (Beispiel MCL)

Systematische Übersicht [Zitat]	Recherchedatum	Unterschied zur Erstellung Referenzstandards	RCT	Vergleichende Kohortenstudien	1-armige Studien
Sancho 2024 [26]	05/2022	< 1 Jahr	nein	ja	ja
NICE 2021 [27]	01/2020	3 Jahre, 3 Monate	ja	ja	ja
Wu 2024 [28]	10/2022	< 1 Jahr	ja	nein	ja
MCL: Mantelzell-Lymphom; RCT: randomisierte kontrollierte Studie					

Die 2 ausgewählten systematischen Übersichten Sancho 2024 und Wu 2024 decken die gesamte Fragestellung ab, sind jedoch breiter als die eigentliche Fragestellung. Da die Publikation NICE 2021 sehr nah an der eigentlichen Fragestellung ist, wurde diese zusätzlich ausgewertet, auch wenn das Recherchedatum weniger aktuell ist.

Identifizierte Primärpublikationen

Abbildung 4 zeigt das Ergebnis der Sichtung der eingeschlossenen Dokumente in den 3 ausgewählten systematischen Übersichten gemäß den Kriterien für den Studieneinschluss. Die Referenzen der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Anhang C.2.4.

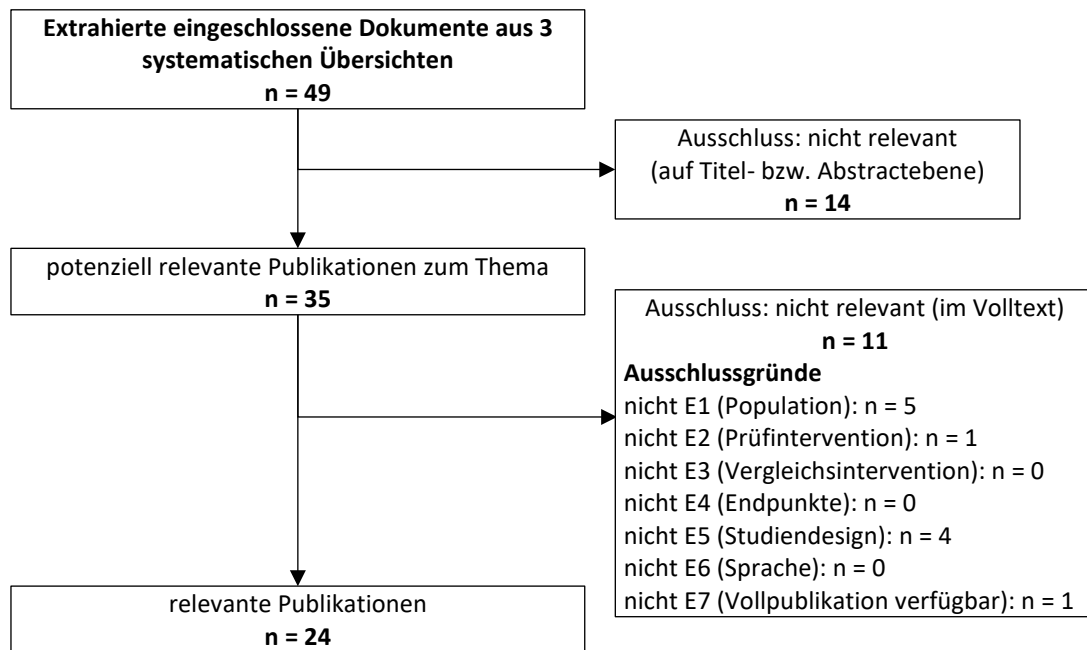


Abbildung 4: Ergebnis der Sichtung der eingeschlossenen Publikationen (Beispiel MCL)

Aufgrund der überschaubaren Anzahl an Publikationen wurde kein Sättigungskriterium angewendet.

Studienpool

Insgesamt wurden aus den 3 systematischen Übersichten 24 Primärpublikationen zur Extraktion von Confounder-Variablen eingeschlossen, davon 1 RCT, 1 Kohortenstudie und 22 1-armige Studien. Die Liste der eingeschlossenen Primärpublikationen aus den systematischen Übersichten findet sich in Anhang C.2.3.

5.1.1.2.2 Ergebnisse der Auswertung

Aus den 24 Primärpublikationen aus den systematischen Übersichten wurden insgesamt 183 Confounder-Variablen extrahiert.

Abgleich mit Referenzstandard

Die identifizierten Variablen wurden mit einem Referenzstandard abgeglichen. Als Referenzstandard wurde eine durch den GBA genannte konkrete Liste an Faktoren herangezogen, die als relevante Confounder für die AbD beim Wirkstoff Brexucabtagen-Autoleucel [14] berücksichtigt werden sollen. Die Ergebnisse des Abgleichs finden sich in Tabelle 10.

Tabelle 10: Abgleich mit Referenzstandard (Beispiel MCL)

Faktoren, die als relevante Confounder für die AbD definiert wurden	Identifiziert durch Primärpublikationen aus SÜ
Alter	ja
Geschlecht	ja
ECOG-Status	ja
Komorbidität	ja
Mantle Cell Lymphoma International Prognostic Index (MIPI) score	ja
Laktatdehydrogenase (LDH)	ja
Leukozytenzahl	ja
Krankheitsstadium	ja
extranodale Erkrankung	ja
Befall des Knochenmarks	ja
Morphologie	ja
B-Symptome	ja
Ki-67	ja
TP53-Mutation	ja
Anzahl vorheriger Therapielinien	ja
vorherige autologe Stammzelltransplantation	ja
Dauer der vorherigen BTK-Inhibitor-Behandlung	ja
Ansprechen auf die vorherige BTK-Inhibitor-Behandlung	ja
BTK: Bruton-Tyrosinkinase; ECOG: Eastern Cooperative Oncology Group; Ki-67: Antigen Ki(Kiel)-67; MCL: Mantelzell-Lymphom; SÜ: systematische Übersicht; TP53: Tumorprotein p53	

Es wurden alle potenziellen Confounder durch die Primärpublikationen aus den systematischen Übersichten identifiziert.

Studientyp

Anhand der Ergebnisse wurde geprüft, welchen Einfluss die Auswahl von bestimmten Studientypen als Quelle auf die Identifikation von potenziellen Confounder-Variablen hat, also ob andere Studientypen neben Kohortenstudien für die Identifikation wichtig sind. Dazu wurden die Primärpublikationen unterteilt in Kohortenstudien, die Adjustierungsfaktoren berichten, und in andere Studientypen (RCTs und 1-armige Studien) bei denen die Patientencharakteristika als Extraktionsquelle genutzt wurden. Tabelle 11 zeigt die Ergebnisse der Auswertung.

Tabelle 11: Anzahl der identifizierten potenziellen Confounder in Abhängigkeit des Studientyps bzw. der Extraktionsquelle^a (Beispiel MCL)

Potenzieller Confounder identifiziert?	Adjustierungsfaktoren aus Kohortenstudien (n = 1)	Patientencharakteristika aus RCT (n = 1), 1-armigen Studien (n = 22)
ja	1	18
unklar	0	0
nein	17	0
a. In den Kohortenstudien wurden die Confounder-Variablen aus den berichteten Adjustierungsfaktoren extrahiert. In den RCTs und den 1-armigen Studien wurden die Patientencharakteristika extrahiert. MCL: Mantelzell-Lymphom; RCT: randomisierte kontrollierte Studie		

Durch das alleinige Betrachten der Adjustierungsfaktoren aus Kohortenstudien wird 1 der 18 potenziellen Confounder identifiziert. Diese Auswertung beruht jedoch nur auf 1 Kohortenstudie, in der lediglich 2 Adjustierungsvariablen genannt wurden [29].

Durch das Betrachten der Patientencharakteristika in den RCTs und den 1-armigen Studien werden alle 18 potenziellen Confounder identifiziert.

Zusammenfassend zeigt sich beim Beispiel MCL, dass durch die Beschränkung auf Kohortenstudien ein gravierender Informationsverlust entsteht.

Aktualität der Recherche

Anhand der Ergebnisse wurde geprüft, welchen Einfluss die Aktualität der Publikationen auf die Anzahl der identifizierten potenziellen Confounder hat. Dazu wurden von den eingeschlossenen Primärpublikationen sukzessive die aktuellsten Publikationen aus der Auswertung herausgelassen. Tabelle 12 zeigt die Ergebnisse der Auswertung.

Tabelle 12: Anzahl der identifizierten potenziellen Confounder in Abhängigkeit der Aktualität (Beispiel MCL)

Potenzieller Confounder identifiziert?	alle Publikationen bis einschließlich 2023 ^a (n = 24)	Publikationen bis einschließlich 2022 (n = 21)	Publikationen bis einschließlich 2021 (n = 18)	Publikationen bis einschließlich 2020 (n = 15)
ja	18	18	18	18
unklar	0	0	0	0
nein	0	0	0	0
a. Das aktuellste Recherchedatum in den systematischen Übersichten ist das Jahr 2022. Aufgrund einer Handsuche in den systematischen Übersichten sind dennoch Publikationen von 2023 enthalten. MCL: Mantelzell-Lymphom; SÜ: systematische Übersicht				

Die Recherche nach systematischen Übersichten erfolgte 08/2024; die Veröffentlichung des Referenzstandards erfolgte 03/2023. Durch den Ausschluss der Publikationen aus dem Jahr 2023, 2022 und 2021 ändert sich die Anzahl der identifizierten potenziellen Confounder nicht.

Zeitlicher Aufwand

Ein Vergleich des zeitlichen Aufwands ist beim Beispiel MCL nicht möglich, da keine bekannten Daten zur Literaturrecherche bei einem Verfahren zur Confounderidentifikation über Primärstudien vorliegen. Deshalb wurde für dieses Beispiel auf eine Betrachtung des zeitlichen Aufwandes verzichtet.

5.1.1.3 Zusammenfassung Confounderidentifikation über systematische Übersichten und die darin enthaltenen Primärstudien

Anhand von 2 Anwendungsbeispielen wurde geprüft, ob als Weiterentwicklung des Ansatzes von Pufulete 2022 eine Aufwandsreduktion bei der Informationsbeschaffung möglich ist. Konkret wurde geprüft, ob ein Rückgriff auf systematische Übersichten zur Identifikation von Primärpublikationen einen ausreichend sensitiven Ansatz zur Confounderidentifikation darstellt. Aus den Ergebnissen kann abgeleitet werden, dass bei Vorliegen aktueller, qualitativ hochwertiger systematischer Übersichten, die die gesamte Fragestellung abdecken, über die darin eingeschlossenen Studien ein Großteil der potenziell relevanten Confounder identifiziert werden kann. Dabei wird, wie bereits in GA23-02 beschrieben, vorausgesetzt, dass bei der Informationsbeschaffung zusätzlich auch Experteninterviews bzw. klinische Expertise einbezogen werden.

Allerdings wurden nur 2 Themen ausgewertet und die Aussagekraft ist daher eingeschränkt.

Einfluss Studientyp

Zudem bestand die Frage, welchen Einfluss die Auswahl von bestimmten Studientypen als Quelle für die Identifikation von potenziellen Confoundern hat.

Beide Beispiele zeigen, dass der Informationsverlust am geringsten war, wenn alle verfügbaren Studientypen bzw. Quellen genutzt werden. Am Beispiel RRMS waren dies RCTs und Kohortenstudien; beim Beispiel MCL auch 1-armige Studien, da dort nur wenige Studien zu anderen Studientypen vorlagen.

Bedeutung der Aktualität der Recherche

Eine weitere Frage stellte sich hinsichtlich der benötigten Aktualität der Recherche. Insgesamt zeigte sich über den zeitlichen Verlauf keine gleichmäßige Akkumulation bzw. Sättigung der identifizierten potenziellen Confounder. Bei beiden Beispielen war das Heranziehen von Primärstudien mit einem Erscheinungsjahr bis 1 oder 2 Jahre vor der Erstellung des Referenzstandards notwendig, um nahezu alle potenziellen Confounder identifizieren zu

können. Allerdings kommt es durch das Recherchedatum in den ausgewerteten systematischen Übersichten zu einem weiteren zeitlichen Abstand von bis zu 2 Jahren (laut AMSTAR-2 sollte dies für eine hochwertige Übersicht nicht älter als 24 Monate sein [30]). Aus diesem Grund sollte die Recherche bei einem prospektiven Studiendesign zu Beginn der Datenerhebung – bzw. bei einem retrospektiven Studiendesign zum Zeitpunkt der Finalisierung des Studienprotokolls – nicht älter als 12 Monate sein. Insbesondere bei einem prospektiven Studiendesign ist eine entsprechende Aktualität unbedingt erforderlich, da die Analyse der Ergebnisse nicht zum Zeitpunkt der Confounderidentifikation, sondern ggf. erst mehrere Jahre später (nach Abschluss der Datenerhebung) durchgeführt wird. Es ist jedoch zu beachten, dass es ggf. Themen geben kann, bei denen die Recherche jünger als 12 Monate sein muss. Dies sollte themenspezifisch geprüft werden.

Aufwandsreduktion projektspezifisch möglich

Es konnte anhand des RRMS-Beispiels gezeigt werden, dass die Informationsbeschaffung über systematische Übersichten im Vergleich zum Vorgehen der externen Sachverständigen [8] grundsätzlich zeitsparender und eine Möglichkeit zur deutlichen Aufwandsreduktion sein kann. So umfasste der Aufwand für die Informationsbeschaffung inklusive der Extraktion ca. 39 Stunden versus 109 Stunden über eine Suche angelehnt an Pufulete 2022.

Hier ist jedoch zu beachten, dass dies nur in Situationen möglich ist, in denen die gesamte Fragestellung über systematische Übersichten abgedeckt werden kann. Beim Beispiel RRMS war dies nicht möglich. Hier sind für die zusätzliche Suche nach Primärstudien weitere 67 Stunden an Arbeitsaufwand geschätzt worden. Sofern eine zusätzliche Suche nach Primärpublikation notwendig ist, entspricht der Aufwand eher dem einer Informationsbeschaffung nach Primärpublikationen (Pufulete 2022).

5.1.2 Einfluss des Sättigungskriteriums auf die Identifikation potenzieller Confounder-Variablen

In GA23-02 [8] wählten die externen Sachverständigen einen sehr sensitiven Ansatz zur Confounderidentifikation: Während alle Publikationen von 2020 bis 2023 zur Extraktion herangezogen wurden (d. h. ohne Anwendung eines Sättigungskriteriums), wurde für Publikationen vor 2020 ein Sättigungskriterium von 15 verwendet, d. h. die Extraktion wurde gestoppt, wenn innerhalb von 15 konsekutiven Publikationen keine neue potenzielle Confounder-Variable identifiziert wurde. Bisher blieb unklar, inwiefern der potenzielle Mehraufwand (im Sinne einer höheren Anzahl an zu extrahierenden Publikationen) im Vergleich zur Verwendung eines restriktiveren Sättigungskriteriums (von z. B. 10, wie in Pufulete 2022 [12]) den potenziellen zusätzlichen Gewinn an Confounder-Variablen rechtfertigt. Das Ziel war es, beispielhaft anhand des Datensatzes aus GA23-02 den Einfluss des Sättigungskriteriums auf die Anzahl identifizierter Confounder-Variablen und auf die für die Extraktion benötigte Studienzahl zu untersuchen.

Insgesamt wurden in GA23-02 160 potenzielle Confounder-Variablen identifiziert. Dabei wurden 125 potenzielle Confounder-Variablen über die Literaturrecherche aus 80 Studien identifiziert, 68 der 125 Variablen aus 40 Publikationen vor 2020 (d. h. unter Verwendung eines Sättigungskriteriums von 15). Weitere 35 Confounder wurden über Experteninterviews genannt. In GA23-02 liegt eine detaillierte Dokumentation (separate Excel-Datei [31]) der pro Studie extrahierten Confounder vor. Dabei entspricht die Nummerierung der Publikationen der Extraktionsreihenfolge. Für die Interpretation der Ergebnisse ist zu beachten, dass es sich bei den 80 eingeschlossenen Publikationen um diejenigen Studien handelt, aus denen die externen Sachverständigen die Confounder-Variablen extrahierten. Der Datensatz stellt somit keine zufällige Stichprobe bestehend aus allen identifizierten Primärpublikationen dar.

Anhand dieses Datensatzes aus GA23-02 (mit n = 80 eingeschlossenen Publikationen) wurden verschiedene Simulationen durchgeführt, in denen die 80 Publikationen 10 000-mal zufällig angeordnet wurden und für die Sättigungskriterien 15, 10 und 5 untersucht wurde, wie viele der insgesamt 125 Confounder identifiziert werden und welche Anzahl an Publikationen dafür ausgewertet werden müssen. Das Ergebnis der Simulation ist in Tabelle 13 dargestellt.

Tabelle 13: Sättigungskriterium – Ergebnis der Simulation einer zufälligen Anordnung (10 000-mal) der 80 Publikationen bei Anwendung eines Sättigungskriteriums von 15, 10 oder 5

Sättigungs- kriterium	MW	Min.	Max.	Perzentile				
				1 %	5 %	25 % (Q1)	50 % (Median)	75 % (Q3)
Anzahl identifizierter Confounder-Variablen								
15	125	93	125	125	125	125	125	125
10	124	47	125	99	124	125	125	125
5	106	18	125	51	67	93	111	125
Anzahl ausgewerteter Publikationen								
15	80	58	80	80	80	80	80	80
10	79	28	80	58	78	80	80	80
5	62	11	80	23	33	50	64	79
Max.: Maximum; Min.: Minimum; MW: Mittelwert; Q1: 1. Quartil; Q3: 3. Quartil								

Im vorliegenden Datensatz führt ein Sättigungskriterium von 15 in > 99 % der Fälle zur Identifikation aller 125 Confounder-Variablen. Bei einem Sättigungskriterium von 10 werden mit einer Wahrscheinlichkeit von min. 75 % ebenfalls alle 125 potenziellen Confounder-Variablen identifiziert. In min. 5 % der Fälle kommt es zu einem Verlust von 1 Confounder-Variable. Die Reduktion der zu berücksichtigenden Publikationen ist bei einem Sättigungskriterium von 10 gering (in min. 95 % der Fälle müssen nur 2 Publikationen weniger ausgewertet werden). Dagegen werden bei einem Sättigungskriterium von 5 mit einer Wahrscheinlichkeit von min. 95 % nur 67 der 125 (54 %) potenziellen Confounder-Variablen

identifiziert und führen so zu einem Verlust von 46 % der Variablen. Die Anzahl der zu berücksichtigenden Publikationen wird im Median auf 64 der 80 (80 %) reduziert.

Die Verteilung (Min- und Max-Werte in Tabelle 13) zeigt, dass die Extraktionsreihenfolge bei der Anwendung eines kleinen Sättigungskriteriums (5) großen Einfluss auf die Anzahl identifizierter Confounder-Variablen sowie die Zahl berücksichtigter Publikationen hat. Der Einfluss ist umso geringer, je größer das Sättigungskriterium ist. Aufgrund dieser Ergebnisse wurde des Weiteren untersucht, zu welchen Ergebnissen eine inhaltlich sinnvolle Reihenfolge (wie z. B. nach Publikationsdatum) führt. Die Ergebnisse nach Publikationsdatum (aufsteigend und absteigend) sind in Anhang D dargestellt, lassen aber aufgrund des beispielhaften Charakters keine übergeordnete Schlussfolgerung zu.

Insgesamt ist bei der Interpretation der vorliegenden Ergebnisse zu beachten, dass es sich lediglich um 1 Beispiel handelt und in diesem nur für die Publikationen nach 2020 alle identifizierten Primärpublikationen enthalten waren. Der Verlust einzelner Confounder-Variablen muss nicht automatisch mit dem Verlust eines Confounders einhergehen, da ggf. mehrere der identifizierten Confounder-Variablen im nachfolgenden Schritt zu Confoundern zusammengeführt oder basierend auf der Fragestellung ausgeschlossen werden. In diesem Beispiel wäre durch die Verwendung eines Sättigungskriteriums von 10 (anstelle von 15 bzw. keinem) in min. 75 % der Fälle dieselbe Anzahl an Confoundern identifiziert worden. Dieser potenzielle Informationsverlust bei einem Sättigungskriterium von 10 wird als vertretbar eingeschätzt. Allerdings wäre der Aufwand in höchstens 25 % der Fälle um 2 Publikationen reduziert worden. Dagegen wird bei einem Sättigungskriterium von 5 der Informationsverlust als zu hoch eingeschätzt.

Insgesamt unterstützt das vorliegende Beispiel das von Pufulete 2022 vorgeschlagene Sättigungskriterium von 10.

Einfluss des Studientyps (RCT oder Beobachtungsstudie)

In den oben betrachteten Datensatz gehen alle Publikationen unabhängig vom Design der jeweiligen Studie ein (RCT und Beobachtungsstudien). Im Unterschied zu Beobachtungsstudien, die Adjustierungsfaktoren berichten, dienen in RCTs die dargestellten Patientencharakteristika (üblicherweise in Tabelle 1 der Publikation) als Extraktionsquelle für potenzielle Confounder-Variablen. Daher kann z. B. angenommen werden, dass in RCTs häufig dieselben Charakteristika abgebildet werden und eine Sättigung somit schneller erreicht wird. Auch in einer Stellungnahme wurde vorgeschlagen, das Sättigungskriterium für RCTs und 1-armige Studien zu reduzieren (bspw. auf einen Wert von 5). Um zu untersuchen, ob es bei RCTs zu einer schnelleren Sättigung der identifizierten Confounder-Variablen kommt, wurden die 27 in GA23-02 eingeschlossenen RCTs gesondert betrachtet und analog zum oben

beschriebenen Vorgehen 10 000-mal zufällig angeordnet. Das Ergebnis ist in Tabelle 14 dargestellt.

Tabelle 14: Sättigungskriterium – Ergebnis der Simulation einer zufälligen Anordnung (10 000-mal) der 27 RCTs bei Anwendung eines Sättigungskriteriums von 15, 10 oder 5

Sättigungs- kriterium	MW	Min.	Max.	Perzentile				
				1 %	5 %	25 % (Q1)	50 % (Median)	75 % (Q3)
Anzahl identifizierter Confounder-Variablen								
15	40	40	40	40	40	40	40	40
10	40	35	40	40	40	40	40	40
5	39	16	40	29	35	40	40	40
Anzahl ausgewerteter Publikationen								
15	27	27	27	27	27	27	27	27
10	27	22	27	27	27	27	27	27
5	26	8	27	14	19	27	27	27
Max.: Maximum; Min.: Minimum; MW: Mittelwert; Q1: 1. Quartil; Q3: 3. Quartil; RCT: randomisiert kontrollierte Studie								

Bei einem Sättigungskriterium von 10 werden mit einer Wahrscheinlichkeit von min. 99 % alle 40 potenziellen Confounder-Variablen identifiziert. Die Anzahl auszuwertender Publikationen ändert sich nicht. Bei einem Sättigungskriterium von 5 werden in min. 5 % der Fälle 5 der 40 (12,5 %) potenziellen Confounder-Variablen nicht identifiziert. Die Anzahl der zu berücksichtigenden Publikationen reduziert sich gleichzeitig auf 19 der 27. Die alleinige Betrachtung von RCTs führt somit im vorliegenden Datensatz nicht zu einer schnelleren Sättigung, da zur Identifikation aller 40 Confounder-Variablen jeweils 27 Publikationen notwendig waren.

Basierend auf den Ergebnissen des vorliegenden Datensatzes wird die Anwendung verschiedener Sättigungskriterien je nach Studiendesign nicht als zielführend eingeschätzt. Zusätzlich müsste die Selektion in 2 Schritten erfolgen, in denen die Publikationen zunächst nach ihrem Studiendesign aufgeteilt würden, um anschließend die unterschiedlichen Sättigungskriterien anzuwenden. Vor dem Hintergrund einer Aufwandsreduktion wird eine gemeinsame Betrachtung aller Publikationen unabhängig vom Studiendesign empfohlen.

5.2 Aufwandsreduktion bei der Datenerhebung

5.2.1 Systematische Ansätze zur Reduktion eines Confounder-Sets

Ergebnis der orientierenden Informationsbeschaffung

Das Ergebnis der orientierenden Recherche nach Publikationen zu systematischen Ansätzen zur Reduktion eines Confounder-Sets, basierend auf inhaltlichen Überlegungen und

unabhängig vom auszuwertenden Datensatz (Kriterien für den Studieneinschluss siehe Abschnitt 4.2.1), wird nachfolgend dargestellt. In der Anhörung zum vorläufigen Rapid Report A25-13 wurde bereits eine Vielzahl von Anregungen genannt. Deshalb wurde der Fokus dieser Recherche insbesondere auf Ansätze gelegt, die über diese Anregungen hinausgehen. Die Suche fand am 06.11.2025 statt.

Durch die Recherche wurden 3 Publikationen [32-34] identifiziert, die zumindest in Teilen das Thema der Reduktion eines Confounder-Sets behandeln. In 2 der 3 Publikationen [32,33] wurden prognostische Faktoren für den Endpunkt durch eingebundene klinische Expertinnen und Experten entsprechend ihrer Bedeutung kategorisiert und dadurch die 10 wichtigsten Faktoren identifiziert. In der 3. Publikation [34] wurden die mittels systematischer Literaturrecherche identifizierten Faktoren nur weiter berücksichtigt, wenn diese in mindestens 2 Studien berichtet wurden und signifikant mit dem Gesamtüberleben und / oder dem Ansprechen assoziiert waren. Daraufhin wurden die Faktoren ebenfalls von klinischen Expertinnen und Experten (n = 5) entsprechend ihrer prognostischen Bedeutung für die Endpunkte in 5 Gruppen kategorisiert und anhand des konsentierten Rankings die wichtigsten Variablen ausgewählt. Da alle 3 Publikationen ein Ranking der Variablen vornehmen, werden diese in Abschnitt 5.2.1.8 zur Relevanz der Confounder – Bewertung der Stärke des Einflusses auf die Effektschätzung berücksichtigt. Weitere Ansätze zur Selektion von Confoundern, welche über die in der Anhörung bereits genannten Anregungen hinausgehen und nicht auf statistischen Verfahren basieren, wurden nicht gefunden.

Auch die Website-Suchen nach methodischen Dokumenten der Zulassungsbehörden EMA und FDA sowie der HTA-Agenturen lieferten keine Hinweise zum methodischen Vorgehen zur systematischen Reduktion eines Confounder-Sets. Zwar wird z. B. durch die EMA [35] und das NICE [36] darauf hingewiesen, dass die getroffenen Annahmen über kausale Zusammenhänge zwischen Kovariaten, Interventionen und Endpunkten idealerweise mittels DAGs dargestellt werden sollten. Abgesehen von den Hinweisen zur Anwendung von DAGs (inkl. der sachgemäßen Berücksichtigung von *Mediatoren*, *Collidern* und *Instrumentenvariablen*) werden jedoch keine relevanten Vorgaben zur Selektion von Confoundern gemacht (siehe auch Abschnitt 5.2.1.7).

Geprüfte Kriterien

Zum einen wurde anhand des umfangreichen Datensatzes identifizierter potenzieller Confounder-Variablen in der Indikation RRMS geprüft, ob sich Kriterien ableiten lassen, auf deren Grundlage Confounder-Variablen identifiziert werden können, auf deren Erhebung verzichtet werden kann. Zum anderen wurden die verschiedenen, in der Anhörung zum vorläufigen Rapid Report A25-13 [37] vorgebrachten Anregungen dahin gehend geprüft, ob diese geeignete Maßnahmen für eine mögliche Aufwandsreduktion darstellen.

Basierend auf dem RRMS-Datensatz konnten außerhalb der Identifikation von Schwellenwerten hinsichtlich der Nennung in der Literatur oder der Berücksichtigung der Art der Quelle keine weitergehenden Kriterien, als die in der Anhörung zu A25-13 bereits genannten, identifiziert werden. Beide Kriterien (Schwellenwerte und Art der Quelle) wurden ebenfalls in der Anhörung genannt.

Nachfolgend werden zunächst die einzelnen Maßnahmen beschrieben und abschließend geeignete und nicht geeignete Maßnahmen gegenübergestellt. Grundsätzlich ist dabei zu beachten, dass einzelne Maßnahmen ggf. mit einem höheren Aufwand während der Planung der Studie einhergehen können, diese anschließend aber – insbesondere in einem prospektiven Studiendesign – den Aufwand der Erhebung, bspw. durch die Reduktion der Anzahl der in der Studie zugrunde liegenden Datenquelle (z. B. Register) zu erhebenden Variablen, deutlich erleichtern. Insgesamt wird der Gesamtaufwand, bestehend aus Planung und Erhebung, betrachtet. Auch wenn bei den folgenden Ausführungen primär von einem prospektiven Studiendesign ausgegangen wird, gelten die daraus abgeleiteten Empfehlungen nichtsdestotrotz gleichermaßen für retrospektive Studien.

5.2.1.1 Anlegen von Schwellenwerten hinsichtlich der Häufigkeit der Nennung in der Literatur

In der Anhörung zum vorläufigen Rapid Report A25-13 [37] stellte das Anlegen von Schwellenwerten hinsichtlich der Häufigkeit der Nennung in der Literatur zur Reduktion der Zahl zu erhebender Confounder eine Anregung zur Aufwandsreduktion dar. Die zugrunde liegende Annahme ist, dass die Häufigkeit der Nennung mit der Relevanz eines Confounders korreliert. Hierzu sind zum jetzigen Zeitpunkt jedoch keine Untersuchungen bekannt. Dieses Vorgehen, nur Variablen, welche z. B. mindestens in 2 Studien berichtet wurden, weiter zu berücksichtigen, wird teilweise auch in der Literatur [34] angewendet, ohne dass dieses Vorgehen inhaltlich begründet wird [34].

Zu diesem Kriterium wurde basierend auf dem umfangreichen Datensatz aus GA23-02 [8] untersucht, ob sich Kriterien für Schwellenwerte ableiten lassen, auf deren Grundlage Confounder-Variablen selektiert werden können.

Variablen, welche im Interview von klinischen Expertinnen und Experten genannt wurden, gingen in die Untersuchung nicht ein, da diesen ein besonderer Stellenwert beigemessen wird. Die Untersuchung des Einflusses eines möglichen Schwellenwertes wurde anhand der Liste der verdichteten 28 Confounder durchgeführt und anschließend eingeordnet. Das Anlegen einer Schwelle von beispielsweise $n = 1$ Publikationen bedeutet, dass Variablen, welche nur 1-mal in der Literatur genannt wurden (und auch nicht im Interview), nicht weiter berücksichtigt werden. Ein Anlegen von Schwellen auf die Liste aller 160 identifizierten Confounder-Variablen wurde als nicht zielführend erachtet, da das Ergebnis stark abhängig vom Detailgrad

der Extraktion ist (z. B. wurden für die Variable gesundheitsbezogene Lebensqualität 13 verschiedene Instrumente bzw. Subskalen jeweils nur 1-mal aus der Literatur extrahiert).

Das Anlegen einer Schwelle von $n = 1$ Publikationen bei der verdichteten Confounder-Liste führt zum Verlust der Confounder Tabakkonsum und kürzliche Schwangerschaft (in den letzten 3 Monaten). Beide Confounder wurden aus Beobachtungsstudien extrahiert. Eine Schwelle von $n = 2$ Publikationen führt zusätzlich zum Verlust der Confounder Bildungsgrad und Fatigue.

Die Frage, inwieweit die einmalige Nennung in der Literatur bedeutet, dass der potenzielle Confounder verzichtbar ist, kann letztlich nicht eingeschätzt werden, da es eine Frage der Kausalität ist, ob eine bestimmte Variable einen Confounder darstellt.

Insgesamt wird die Anwendung einer Schwelle nicht empfohlen. Da bereits durch das Anlegen eines Sättigungskriteriums (siehe Abschnitt 5.1.2) eine gewisse Unsicherheit in Kauf genommen wird, wird nicht empfohlen weitere potenzielle Confounder allein auf Basis einer Schwelle auszuschließen und damit die in Kauf genommene Unsicherheit zu vergrößern. Eine inhaltliche Rationale für die Setzung einer (arbiträren) Schwelle konnte nicht identifiziert werden. So hängt das Anlegen einer Schwelle stark von der Anzahl identifizierter Studien ab. Falls nur wenige Studien im Anwendungsgebiet vorliegen, sollte keine zusätzliche Schwelle verwendet werden. Insbesondere in Indikationen mit absehbar wenig umfassender Studienlage (z. B. bei seltenen Erkrankungen) ist die Wahrscheinlichkeit höher, dass potenzielle Confounder-Variablen nur einmal in der Literatur vorkommen. Zudem wird durch die Verwendung eines Sättigungskriteriums die Wahrscheinlichkeit erhöht, Confounder-Variablen ggf. nur einmal in der Literatur zu finden. Zusätzlich kann es sein, dass ein Confounder erst kürzlich identifiziert wurde und deshalb nur in einer der aktuellsten Studien berichtet wird.

5.2.1.2 Berücksichtigung der Art der Quelle

Bereits in Abschnitt 5.1.1.1.2 wurde die Berücksichtigung der Art der Quelle, also der Einfluss des Studientyps (RCT vs. nicht randomisierte Beobachtungsstudie) auf die Confounderidentifikation untersucht. Hintergrund für diese Untersuchung ist die Annahme, dass die aus nicht randomisierten Studien extrahierten Adjustierungsvariablen zielgerichtet als Confounder ausgewählt wurden, und damit ggf. spezifischer sind als die in RCT beschriebenen Patientencharakteristika (in der Regel in Tabelle 1 einer Publikation). Aufgrund dessen wurde untersucht, ob die Berücksichtigung von RCT – und damit der in Tabelle 1 genannten Patientencharakteristika – zu einem relevanten Informationsgewinn für die Confounderidentifikation führt.

In Abschnitt 5.1, der übergeordnet die Frage untersucht, inwieweit der Rückgriff auf systematische Übersichten und die darin enthaltenen Primärpublikationen ausreichend ist,

dienten entsprechend die identifizierten systematischen Übersichten und die darin enthaltenen Primärstudien als Datengrundlage. Im Gegensatz dazu bedienen sich die im Folgenden beschriebenen Untersuchungen dem RRMS-Datensatz, der ausschließlich auf einer Suche nach Primärstudien basiert.

Insgesamt wurden im vorliegenden Datensatz 13 der 160 Variablen ausschließlich über RCTs identifiziert. Von diesen wurden im 2. Schritt der Zusammenführung 10 zu verschiedenen Confoundern verdichtet und 2 basierend auf der Fragestellung der Studie CONFIRM ausgeschlossen. Letztlich wurde in GA23-02 nur der Confounder Volumen T1-hypointenser Läsionen ausschließlich über 7 Publikationen zu RCTs identifiziert. Alle anderen Confounder wurden mindestens über Beobachtungsstudien oder Interviews identifiziert.

Im Workshop zur DAG-Erstellung (siehe Abschnitt 5.2.1.7) haben die eingebundenen Klinikerinnen und Kliniker im Konsens zurückgemeldet, dass die Variable Volumen T1-hypointenser Läsionen im Versorgungsalltag nicht erhoben wird. Aus diesem Grund wurde der Confounder im Projektverlauf zusammen mit dem Volumen der T2-Läsionen zum Confounder Volumen der Läsionen, Veränderung im Vorjahr verdichtet (siehe auch Tabelle 27). Darin enthaltene Variablen, wie z. B. Läsionslast und Größe der Läsionen, wurde unter anderem auch im Interview genannt.

Grundsätzlich ist es plausibel, dass in interventionellen Studien (RCTs) auch Variablen erhoben und berichtet werden, welche im Versorgungsalltag nicht gemessen und berücksichtigt werden. Im vorliegenden Beispiel wurden – unter Berücksichtigung der Informationen aus dem Workshop – keine zusätzlichen Confounder über RCTs identifiziert und die Berücksichtigung von RCTs wäre verzichtbar gewesen. Da es sich hierbei lediglich um 1 Beispiel handelt, ist dieses Ergebnis nicht ausreichend, um generell RCTs als Quelle für die Confounderidentifikation auszuschließen. Letztlich wird die Relevanz von RCTs stark von der Datenlage in der jeweiligen Indikation abhängig sein. Wenn es eine relativ breite Datenbasis von Beobachtungsstudien gibt (dies ist nach dem Screening bekannt), kann unter Umständen auf die Extraktion aus RCTs verzichtet werden. Dieses Vorgehen sollte jedoch a priori definiert werden. In Hinblick auf Fragestellungen in Indikationen mit absehbar weniger umfassender Studienlage (z. B. bei seltenen Erkrankungen) ist es nicht zielführend Studien, allein auf Basis des Studiendesigns auszuschließen. Hier kann beim Fehlen von geeigneten RCTs und Kohortenstudien auch angezeigt sein, auf 1-armige Studien (inklusive Vorher-Nachher-Vergleichen) und die darin in tabellarischer Form aufgeführten Patientencharakteristika zurückzugreifen. Genauso besteht die Möglichkeit in diesen Fällen z. B. Leitlinien in die Recherche einzubeziehen (siehe auch GA23-02 [8]).

5.2.1.3 Berücksichtigung der Qualität der Quelle

Eine weitere Anregung war die Qualität einer Quelle, aus der Confounder extrahiert werden, zu berücksichtigen.

Die Qualität einer Quelle wird von den Stellungnehmenden bzw. in der Literatur (wie z. B. in [32-34]) im Kontext der Confounderidentifikation durch unterschiedliche Kriterien beschrieben, wie z. B.:

- Vorhandensein einer Beschreibung der Methodik, wie die Confounder / Variablen identifiziert wurden (Berichtsqualität),
- Vorab-Registrierung der Studie und / oder a priori Definition der Confounder,
- Risk of Bias (Assessment),
- Größe der Studie.

Zusätzlich zu diesen Kriterien wurde auch angeregt, das Studiendesign zur Abwägung der potenziellen Relevanz eines Confounders zu berücksichtigen. Dieses Kriterium wurde bereits in den vorherigen Abschnitten 5.1.1.1.2 und 5.2.1.2 beschrieben.

Das Kriterium „Vorhandensein einer Beschreibung der Methodik, wie die Confounder / Variablen identifiziert wurden“ wurde beispielhaft für einige Variablen anhand des RRMS-Datensatzes aus GA23-02 geprüft. Für die meisten der betrachteten Publikationen lagen keine Angaben zur Confounderidentifikation vor, teilweise wurde auf „Subject-matter-Knowledge“ verwiesen. Auch in GA23-02 wurde von den externen Sachverständigen festgestellt, dass die Schlagwort-Suche (z. B. nach „causal“ oder „dag“) in den untersuchten Publikationen wenig erfolgreich war. Bei den vorliegenden Publikationen würde die Anwendung dieses Qualitätskriteriums somit zum Ausschluss der meisten Studien führen. In der Publikation Karim 2022 wurde eine ähnliche Beobachtung berichtet: Alle Studien gaben die Liste der für das Propensity-Score-Modell verwendeten Variablen an, doch nur 21 % der Studien erwähnten, wie diese Variablen ausgewählt wurden [38]. Die Beobachtung, dass zumeist eine detaillierte Beschreibung der Methodik der Confounderidentifikation fehlt, wurde ebenfalls in einer aktuellen Publikation von Neumann 2026 [39] beschrieben. Einen Ausschluss aller Publikationen zu erwägen, welche die Methodik der Confounderidentifikation nicht hinreichend beschreiben, würde insbesondere in Hinblick auf Fragestellungen in Indikationen mit absehbar wenig umfassender Studienlage (z. B. bei seltenen Erkrankungen) dazu führen, dass die ohnehin schon eingeschränkte Datengrundlage für die Confounderidentifikation nahezu wegfallen würde. Die Anwendung dieses Kriteriums wird daher zum jetzigen Zeitpunkt als nicht zielführend eingeschätzt. Sollte in Zukunft eine systematische Confounderidentifikation für eine Fragestellung identifiziert werden, die als hinreichend anwendbar

eingeschätzt wird und deren Methodik hinreichend beschrieben ist, könnten die dort identifizierten Confounder verwendet und die Literaturrecherche beendet werden.

Gleiches gilt für das Kriterium der Vorab-Registrierung einer Studie. Hier müsste im Registereintrag jeweils das Studienprotokoll verfügbar sein, um zu prüfen, ob die Adjustierungsvariablen genannt werden und darüber hinaus die Methodik der Confounderidentifikation beschrieben wird. Eine stichprobenartige Analyse der im vorliegenden RRMS-Datensatz eingeschlossenen Publikationen zeigte, dass für den Großteil keine Identifikationsnummer gefunden wurde. Analog zum ersten Kriterium würde die Prüfung, ob eine Studie vorab registriert wurde, mit großem Aufwand einhergehen und auch hier zum Wegfall des überwiegenden Anteils an Publikationen führen. Somit wird die Anwendung im Kontext der Confounderidentifikation und -selektion ebenfalls nicht als zielführend eingeschätzt wird.

Des Weiteren wurde zur Beurteilung der Qualität der Quelle eine Risk-of-Bias-Bewertung beschrieben. Die Einschätzung von Verzerrungsaspekten ist ein wesentlicher Gesichtspunkt bei der Interpretation von Studienergebnissen vergleichender Studien [40]. Dabei spielen die Fragestellung, die Art, das Design und die Durchführung der Studien sowie die Verfügbarkeit der Informationen eine Rolle, um systematische Verzerrungen zu identifizieren [40]. Die Berücksichtigung von Confoundern ist u. a. auch ein Bestandteil des ROBINS-I-V2-Tools [41] zur Bewertung der Verzerrung in nicht randomisierten Studien. Eine mögliche Verzerrung von Studienergebnissen ist jedoch für die Selektion der Confounder von untergeordneter Bedeutung. Somit rechtfertigt das Vorliegen einer Verzerrung nicht den Ausschluss einer Variablen.

Ein weiterer Vorschlag sieht vor, die Größe der jeweiligen Studie in die Beurteilung der Qualität der Quelle einfließen zu lassen. Die Unterscheidung, ob der Confounder auf Grundlage von wenigen Patientinnen und Patienten oder einer größeren Fallzahl untersucht wurde, kann bei der Extraktion statistisch signifikanter Effekte von Relevanz sein. Allerdings wurde in GA23-02 – wie bereits zuvor erwähnt – vorgeschlagen, alle Variablen unabhängig von ihrer statistischen Signifikanz zu extrahieren. Daher hat die Größe der Studie für das in GA23-02 vorgeschlagene Vorgehen keine Bedeutung.

Insgesamt ist zu beachten, dass die Qualität der Quelle vor dem Hintergrund der Confounderidentifikation zu betrachten ist. Von den oben angeführten Kriterien ist das 1. Kriterium (Vorhandensein einer Beschreibung der Methodik, wie die Confounder / Variablen identifiziert wurden) der wesentliche Aspekt, der die Qualität der Confounderidentifikation definiert. Allerdings wird dieses Kriterium aufgrund der zum jetzigen Zeitpunkt in der Regel unzureichenden Berichterstattung zur Confounderidentifikation aktuell nicht als zielführend erachtet. Alle anderen genannten Aspekte (Vorab-Registrierung der Studie und / oder a priori Definition der Confounder, Risk of Bias [Assessment], Größe der Studie) werden im Kontext

der Confounderidentifikation und -selektion per se nicht als zielführende Maßnahme eingeschätzt.

5.2.1.4 Berücksichtigung von Korrelationen

Weiterhin wurde angeregt, die Korrelation von Variablen bei der Auswahl der zu erhebenden potenziellen Confounder stärker zu berücksichtigen. Über die Korrelation wird ausgedrückt, inwieweit 2 Variablen miteinander zusammenhängen. Eine Korrelation beschreibt nur einen Zusammenhang, der nicht gleichzusetzen ist mit einem kausalen Zusammenhang. Auch potenzielle Confounder-Variablen, welche zunächst nicht das gleiche Konstrukt wie beispielsweise Lebensqualität und Komorbiditäten messen, können miteinander korreliert sein.

Ziel für die Confounderidentifikation wäre es, durch Berücksichtigung dieser Korrelationen eine weitere Reduktion von Variablen zu erreichen. Der zugrunde liegende methodische Ansatz ist in Abbildung 5A grafisch dargestellt: unter der Annahme, dass es einen kausalen Pfad von Confounder C1 zum Confounder C2 gibt, sind die Confounder auch miteinander korreliert. Es würde in diesem einfachen Beispiel ausreichen, einen der beiden Confounder für die Adjustierung zu berücksichtigen, dabei würde in der Regel die Auswahl von C2 zu einer präziseren Schätzung des Endpunkts führen. Dabei ist allerdings zu beachten, dass eine Korrelation nicht unbedingt auch einen kausalen Zusammenhang bedeutet. Des Weiteren könnte mit der Auswahl nur eines Confounders die *Multikollinearität* sowie Konvergenzprobleme insbesondere bei binären Confoundern reduziert werden. Publikationen, in denen dieser Ansatz zur Reduktion eines Confounder-Sets herangezogen wurde, wurden allerdings nicht identifiziert (weder in den Stellungnahmen zu A25-13 genannt, noch über die orientierende Informationsbeschaffung identifiziert). Nicht berücksichtigt wird dabei außerdem, ob stark miteinander korrelierte Variablen nicht zusätzliche, unabhängige Informationen bezüglich des Confounding enthalten können, sodass trotz Korrelation beide Variablen für eine erfolgreiche Confounderadjustierung benötigt werden (siehe Szenario B in Abbildung 5).

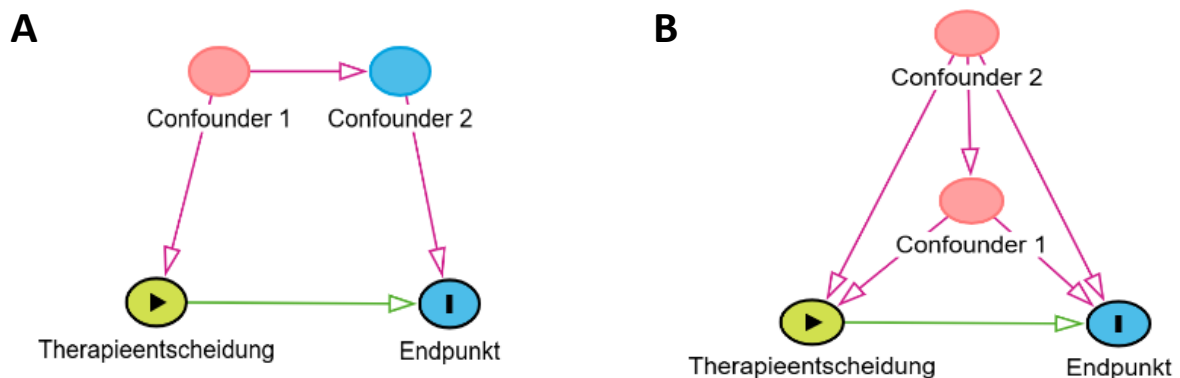


Abbildung 5: Beispielhafte Konstellationen 2 korrelierter Confounder (1 und 2):

A) Aufgrund der Korrelation der beiden Confounder 1 und 2 ist die Adjustierung für einen der beiden ausreichend. B) Trotz der Korrelation der Confounder 1 und 2 trägt jeder der beiden Confounder zusätzliche Informationen für den Zusammenhang zwischen Therapieentscheidung und Endpunkt bei. Abbildung erstellt mit der Software DAGitty [42,43].

Um korrelierte Variablen zu identifizieren kann entweder (sofern verfügbar) eine Analyse anhand bereits vorhandener Datensätze, z. B. eines Registers, erfolgen (z. B. auf Basis einer Korrelationsmatrix) oder eine literaturbasierte Recherche nach der Korrelation der identifizierten Confounder durchgeführt werden. Hierbei ist zu beachten, dass die Relevanz eines Confounders immer im Zusammenhang mit allen identifizierten Confoundern beurteilt werden sollte. Das bedeutet, dass sowohl für die Analyse retrospektiver Daten als auch bei einer literaturbasierten Recherche idealerweise die Informationen zu allen identifizierten Confoundern aus einer Quelle zur Verfügung stehen.

Für den vorliegenden Bericht wurde dieser korrelationsbasierte Ansatz für die in der Indikation RRMS identifizierten Confounder untersucht. Zunächst wurde versucht, eine fokussierte Recherche nach Publikationen aufzusetzen, welche entsprechende Korrelationen zwischen den einzelnen Confoundern in der Indikation RRMS untersuchen. Dies war jedoch nicht zielführend möglich. Bei den gefundenen Publikationen [44-46] handelte es sich um Zufallsfunde mittels einer unsystematischen Handsuche. Meist wurde die Korrelationsmatrix im jeweiligen Anhang dargestellt und war entsprechend über die systematische Recherche nicht auffindbar. Die über die Handsuche identifizierten Publikationen deckten teils andere Fragestellungen und auch nur einen Bruchteil der identifizierten Confounder ab. Zusätzlich wurde keine Publikation identifiziert, welche gezielt Korrelationen zur Selektion von Confoundern in der Indikation RRMS untersuchte.

Eine datenbasierte Untersuchung der Korrelationen (anhand z. B. einer Korrelationsmatrix) war im vorliegenden Projekt nicht möglich. Dafür wäre ein Datensatz notwendig, der anhand einer für die zu untersuchende Studienpopulation repräsentativen Stichprobe erstellt wurde,

um auf die Koeffizienten zurückgreifen zu können. Nach den bisherigen Erfahrungen aus der AbD lagen in den Datenquellen zum Zeitpunkt der Confounderidentifikation und -selektion nur für einen Teil der identifizierten, potenziell relevanten Confounder retrospektive Daten zur Erstellung einer Korrelationsmatrix vor.

Insgesamt kann die Berücksichtigung von Korrelationen nicht als zielführender systematischer Ansatz zur Reduktion des Confounder-Sets empfohlen werden. Zum einen ist eine systematische Literaturrecherche nach Korrelationen nicht möglich, eine unsystematische Handsuche wiederum sehr aufwendig. In dem anzunehmend sehr seltenen Fall, dass ein Datensatz einer für die zu untersuchende Fragestellung repräsentativen Stichprobe zur Verfügung steht, der mindestens einen Großteil der identifizierten Confounder-Variablen enthält, kann die Untersuchungen von Korrelationen, wie in Abbildung 5A oben skizziert, ein Ansatzpunkt sein, um die Reduktion von Confoundern zu unterstützen. Nicht berücksichtigt wird dabei, ob stark miteinander korrelierte Variablen nicht zusätzliche, unabhängige Informationen bezüglich des Confounding enthalten können, sodass trotz Korrelation beide Variablen für eine erfolgreiche Confounderadjustierung benötigt werden.

5.2.1.5 Fokussierung auf Faktoren, die die Therapieentscheidung beeinflussen

In der Anhörung zu A25-13 wurde des Weiteren vorgeschlagen, sich bei der Identifikation von Confoundern auf Variablen zu beschränken, von denen bekannt ist, dass sie die Therapieentscheidung beeinflussen, ohne dabei zu betrachten, ob sie einen Einfluss auf die Endpunkte haben. Die Idee dabei ist es, möglichst die Randomisierung anhand von Baselinevariablen nachzubilden. Des Weiteren sind dadurch *Mediatoren* und *Collider*, die nach Baseline erhoben werden, ausgeschlossen. Dabei ist zu berücksichtigen, dass bekannte *Instrumentenvariablen*, also solche, die ausschließlich die Therapieentscheidung und nicht den Endpunkt beeinflussen, ausgeschlossen werden sollten (siehe Abschnitt 5.2.1.6). Zur Identifikation von Variablen, die die Therapieentscheidung beeinflussen, können neben der Literaturrecherche und Expertenbefragungen auch Leitlinien – soweit vorhanden – beitragen.

Eine Schwachstelle bei der Konzentration auf Variablen, von denen bekannt ist, dass sie die Therapieentscheidung beeinflussen, liegt in dem Ausschluss solcher Variablen, die „nur“ prognostisch für den Endpunkt sind. Diese Variablen sind keine Confounder und bräuchten für die Propensity-Score-Berechnung nicht berücksichtigt werden, um – in der Theorie – eine unverzerrte Effektschätzung zu ermöglichen. Es könnte allerdings eine Situation auftreten, in denen aus einer prognostischen Variablen ein Confounder wird. Dies kann zum einen der Fall sein, wenn der prognostische Faktor durch Zufall unterschiedlich verteilt ist und damit in der Studienpopulation ein Zusammenhang zwischen dieser Variablen und der Therapieentscheidung vorliegt. Zum anderen können prognostische Faktoren zu Confoundern werden, wenn der prognostische Faktor in den jeweiligen Behandlungsrmen systematisch unterschiedlich verteilt ist. Dies kann zum Beispiel der Fall sein, wenn die Therapiearme aus

unterschiedlichen Datenquellen kommen und sich daher ggf. in der zugrunde liegenden Population unterscheiden. In beiden Situationen führt die Nichtberücksichtigung dieser Variablen zu verzerrten Schätzungen [47-51].

In der Publikation Witte 2019 [47] wurden verschiedene datenbasierte Selektionsmethoden, u. a. die Einschränkung auf Variablen, die nur die Therapieentscheidung beeinflussen, miteinander verglichen. Die Entscheidung, ob eine Variable die Therapieentscheidung beeinflusst, beruhte in ihren Untersuchungen auf statistischen Tests – zwischen Behandlung und Endpunkt – anhand der vorliegenden Studiendaten. Dabei erwies sich dieser Ansatz als der schlechteste bezüglich der Präzision und der Verzerrung. Insbesondere kann hier die Gefahr vorliegen, dass eine Variable, die nur schwach mit der Therapieentscheidung aber stark mit dem Endpunkt zusammenhängt, nicht berücksichtigt wird [51].

Des Weiteren können durch diesen Ansatz Instrumentenvariablen in dem Adjustierungsset enthalten sein, d. h. Variablen die mit sehr großer Wahrscheinlichkeit nur die Behandlung vorhersagen. Die Berücksichtigung dieser Variablen kann sowohl zu verzerrten Effektschätzungen führen als zu einer verringerten Präzision [52] (siehe auch Abschnitt 5.2.1.6).

Zusammenfassend kann die Strategie sich auf diejenigen Variablen zu konzentrieren, die ausschließlich die Therapieentscheidung beeinflussen, nicht empfohlen werden. Von einer Auswahl aufgrund von statistischen Tests ist abzuraten.

5.2.1.6 Verzicht auf instrumentelle Variablen

Von den Stellungnehmenden [37] wurde zudem angeregt, bekannte instrumentelle Variablen nicht als potenzielle Confounder im Adjustierungsset zu berücksichtigen.

Eine Variable ist per Definition eine *Instrumentenvariable*, wenn folgende 3 Bedingungen erfüllt sind: (1) die Variable ist assoziiert mit der Intervention, (2) hat nur über die Intervention Einfluss auf den Endpunkt und (3) die Variable und der Endpunkt haben keine gemeinsamen Ursachen [6]. Wird eine Instrumentenvariable als Adjustierungsfaktor im Modell berücksichtigt, verringert dies im besten Fall nur die Präzision [53]. Bei gleichzeitigem Vorliegen von *unmeasured Confounding* birgt eine Berücksichtigung das Potenzial, zusätzliche Verzerrungen zu verursachen [53-55]. Allerdings ist die Annahme, dass es sich bei einer potenziellen Instrumentenvariable tatsächlich in Beobachtungsstudien um eine solche handelt, in der Praxis nicht überprüfbar [6,53].

Bezogen auf das in GA23-02 vorgeschlagene Vorgehen der Confounderidentifikation ist es durchaus möglich, dass nicht ausschließlich Confounder, sondern ggf. auch andere Variablen wie Instrumentenvariablen identifiziert werden. Da keine allgemeingültige Liste von Instrumentenvariablen existiert, muss für die jeweilige Fragestellung geprüft werden, ob eine

der identifizierten Variablen ggf. die oben genannten Bedingungen erfüllt. Liegt eine „echte“ Instrumentenvariable vor, sollte diese nicht im Adjustierungsset berücksichtigt werden. Da dies immer eine Annahme sein wird, muss letztlich eine potenzielle Verzerrung aufgrund der Adjustierung für eine Instrumentenvariable gegen die Gefahr, wesentliche Confounder aus dem Adjustierungsset auszuschließen, abgewogen werden [56].

Wie Zusammenhänge zwischen Intervention, Endpunkten und den identifizierten Variablen erarbeitet und damit ggf. Instrumentenvariablen ermittelt werden können, ist Bestandteil der Erstellung eines DAG und wird ausführlich im nachfolgenden Abschnitt 5.2.1.7 beschrieben.

5.2.1.7 Erarbeitung gerichteter azyklischer Graphen

In mehreren Stellungnahmen [37] wurde auf gerichtete azyklische Graphen (directed acyclic graphs [DAGs]) als Möglichkeit zur Definition eines Minimalsets an Confoundern und damit zur systematischen Reduktion des zu erhebenden Confounder-Sets hingewiesen. Im Folgenden werden zunächst die methodischen Grundsätze von DAGs erläutert und anschließend die Erarbeitung mithilfe klinischer Expertinnen und Experten beschrieben. Anschließend werden die Ergebnisse dargestellt und abschließend zusammengefasst.

5.2.1.7.1 Methodische Grundsätze

Ein DAG ist ein grafisches Werkzeug, mithilfe dessen die (angenommenen) kausalen Zusammenhänge zwischen Therapieentscheidung, Endpunkt und weiteren Variablen visualisiert werden können. Dafür müssen eine Reihe von Annahmen über Kausalzusammenhänge zwischen den Variablen getroffen werden, welche auf Fachwissen beruhen. DAGs enthalten somit kausale Hypothesen, die z. B. klinische Expertinnen und Experten auf der Grundlage ihres Fachwissens definiert haben [57]. Daher können DAGs auch nur so gut sein, wie das Fachwissen, auf dessen Grundlage sie entwickelt wurden [58]. Durch den DAG werden die getroffenen Annahmen explizit gemacht und für alle transparent dargestellt.

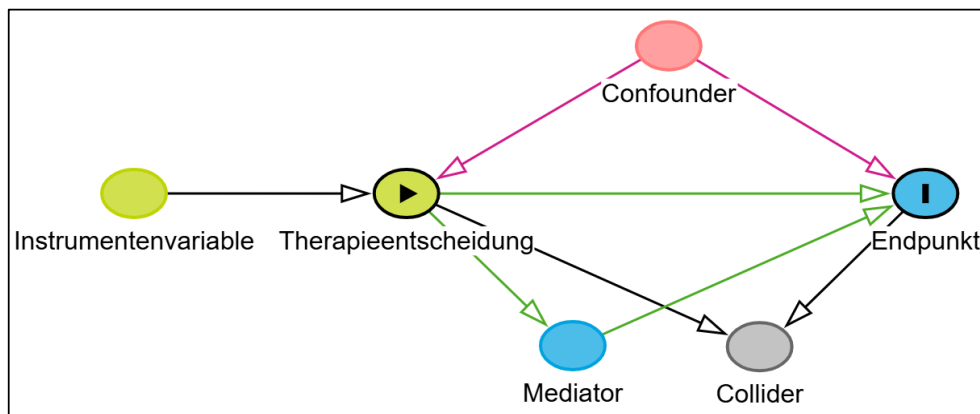
Grundsätzlicher Aufbau eines DAG

Ein DAG besteht aus Knoten und Kanten. Knoten repräsentieren die Variablen einer Studie (wie Therapieentscheidung, Endpunkte und Confounder) und werden durch gerichtete Kanten (Pfeile) miteinander verbunden. Bei der Erstellung eines DAG ist für jedes Variablenpaar zu prüfen, ob für ≥ 1 Patientin oder Patienten ein kausaler Zusammenhang zwischen einer Variablen „X“ und einer anderen Variablen „Y“ besteht oder umgekehrt. Falls ein direkter Zusammenhang besteht, wird ein Pfeil zwischen den Variablen (von X zu Y oder umgekehrt von Y zu X) gezeichnet. Ein Pfeil gibt die Richtung des Zusammenhangs an, beschreibt jedoch weder das Vorzeichen (negativ oder positiv), noch die Stärke oder die Form (z. B. linear oder nicht linear) des Zusammenhangs. In seltenen Fällen kann es notwendig sein, zwischen 2 Variablen bidirektionale Pfeile einzuzeichnen. Dies bedeutet, dass eine nicht gemessene gemeinsame Ursache (*Confounder*) existiert, die beide Variablen beeinflusst. Ein DAG ist somit

eine rein qualitative und keine quantitative Betrachtungsweise. Das Weglassen eines Pfeils impliziert, dass kein kausaler Zusammenhang vorliegt und ist eine wesentlich stärkere statistische Annahme als die, die bei Einschluss eines Pfeils zugrunde liegt [59]. Da keine Variable sich selbst zu einem bestimmten Zeitpunkt und die Zukunft nicht die Vergangenheit bedingen kann, sind DAGs azyklisch aufgebaut und kausale Beziehungen sollten nicht der zeitlichen Richtung widersprechen [59]. Grundsätzlich sollten auch nicht messbare Variablen in ein DAG aufgenommen werden.

Variablentypen

In Abhängigkeit von den Zusammenhängen zur Therapieentscheidung und Endpunkt werden verschiedene Variablentypen unterschieden: *Confounder*, *Collider*, *Mediatoren* und *Instrumentenvariablen* (siehe Abbildung 6). Confounder beeinflussen sowohl die Therapieentscheidung als auch den Endpunkt und führen zu einer Verzerrung des Behandlungseffekts, wenn für sie nicht kontrolliert wird. Ein Collider wird im Gegensatz dazu sowohl von Therapieentscheidung als auch dem Endpunkt beeinflusst. Die Pfeile „kollidieren“ in dieser Variablen. Ein Mediator liegt auf dem direkten Weg zwischen Therapieentscheidung und Endpunkt und eine Instrumentenvariable ist u. a. dadurch definiert, dass sie nur die Therapieentscheidung selbst beeinflusst (siehe auch Abschnitt 5.2.1.5).



Grüne Pfeile stellen kausale Wirkungswege zwischen Therapieentscheidung und Endpunkt dar. Rote Pfeile hingegen Wirkungswege (sogenannte offene Pfade), die den Zusammenhang zwischen Therapieentscheidung und Endpunkt verfälschen. Schwarze Pfeile beeinflussen den Effekt zwischen Therapieentscheidung und Endpunkt nicht.

Abbildung 6: Darstellung der verschiedenen Variablentypen und ihren jeweiligen Zusammenhängen zu Therapieentscheidung und Endpunkt in einem gerichteten azyklischen Graphen (DAG), erstellt mit der Software DAGitty [42,43]

Mithilfe von DAGs können sogenannte Backdoor-Pfade („Hintertür-Pfade“) identifiziert werden. Diese verbinden die Therapieentscheidung mit dem Endpunkt, aber nicht über den direkten kausalen Weg. Pfade, die durch einen oder mehrere Confounder führen, werden als offen bezeichnet und können die Schätzung des Behandlungseffekts verfälschen (rote Pfeile

in Abbildung 6). Daher ist es wichtig für Variablen (Confounder), die auf diesem Weg liegen, zu adjustieren und die offenen Pfade dadurch zu schließen (zu blockieren).

Pfade, auf denen sich ein Collider befindet, sind bereits geschlossen und eine Adjustierung ist nicht notwendig bzw. sogar potenziell schädlich, da diese zu Verzerrungen führen kann. Ebenfalls sollte nicht für Mediatoren adjustiert werden, da damit der kausale Pfad blockiert und ein Teil des interessierenden Effekts eliminiert werden würde. Mithilfe entsprechender Algorithmen kann somit aus dem DAG ermittelt werden, für welche Teilmenge der identifizierten Confounder mindestens adjustiert werden muss, um alle identifizierten Confounder adäquat in der Auswertung zu berücksichtigen. Dies wird auch als *minimal ausreichendes Adjustierungsset* (minimal sufficient adjustment set) bezeichnet [60-63].

In einem DAG können eine Reihe weiterer Variablen bzw. Ereignisse abgebildet werden. Zum Beispiel können konkurrierende Ereignisse, wie Tod krankheitsspezifischer Ursache vs. Tod anderer Ursache [44], dargestellt werden, da diese unter Umständen den Behandlungseffekt verzerren können. Gleiches gilt für die Darstellung potenzieller Selektionsknoten, falls davon auszugehen ist, dass es sich bei der selektierten Population nicht um eine repräsentative Stichprobe der Zielpopulation handelt [64-66]. Knoten für Lost to Follow-up bzw. Drop-out können ebenfalls in ein DAG aufgenommen werden, wenn diese für die zu untersuchende Fragestellung von Bedeutung sind (z. B. für eine Per-Protocol-Analyse).

Weitergehende ausführliche Beschreibungen zur Definition und Erarbeitung von DAGs finden sich z. B. in den Publikationen Digitale 2022 [65], Ferguson 2019 [67], Rodrigues 2022 [68], Tennant 2021 [59] und VanderWeele 2019 [54].

Die Erarbeitung eines DAG und der möglichen Zusammenhänge zwischen verschiedenen Variablen stellt somit eine potenziell geeignete Maßnahme zur Selektion von Confounder dar. Für die vorliegende Ausarbeitung wurde ein DAG auf Basis der in GA23-02 identifizierten 28 Confounder in der Indikation RRMS erarbeitet. Hierzu wurde ein Workshop mit klinischen Expertinnen und Experten durchgeführt. Die angewendete Methodik und die Ergebnisse werden in den folgenden Abschnitten 5.2.1.7.2 und 5.2.1.7.3 beschrieben.

In der Literatur existieren zwar zahlreiche Publikationen zu den methodischen Grundsätzen der DAG-Erstellung, aber nur sehr wenige dazu, wie die praktische Einbindung von Experten(wissen) erfolgen soll [64,67-70]. Aus diesem Grund stand bei der Durchführung des Workshops die Überprüfung der Methode (Proof-of-Concept) im Fokus. Dies umfasst einerseits die Frage, ob die Erarbeitung eines DAG für die vorliegende Fragestellung der Aufwandsreduktion zielführend und machbar ist. Andererseits sollten die gewonnenen Erfahrungen genutzt werden, um praktische Handlungsempfehlungen für die Erstellung eines DAGs sowie die notwendigen zu dokumentierenden Informationen abzuleiten.

5.2.1.7.2 Workshop zur Erstellung eines gerichteten azyklischen Graphen

Die verschiedenen Schritte der Planung und Durchführung des Workshops sind tabellarisch in Tabelle 15 dargestellt und werden in den folgenden Abschnitten textlich erläutert.

Tabelle 15: Schritte der Planung und Durchführung des Workshops zur Erstellung eines DAG

Phase	Umsetzung
geplante Rahmenbedingungen	▪ Einbindung von 3-5 Klinikerinnen und Klinikern ▪ Einbindung 1 Methodikerin mit Expertise in DAGs ▪ Durchführung in Präsenz ▪ Dauer: 1 Tag
Vorbereitung	▪ Vorsortierung der Variablen anhand eines zeitlichen Rasters ▪ Versand der Vorabinformationen für eingebundene Expertinnen und Experten
Durchführung	
Einleitung	▪ Einführung in die Fragestellung ▪ Einführung in DAGs ▪ Besprechung der Variablen und deren Operationalisierung
DAG-Erarbeitung	▪ Ausgangspunkt: Start mit leerem (kantenlosem) DAG ▪ Abfrage des Einflusses der Variablen auf die Therapieentscheidung / die Endpunkte ▫ (bei Weglassen eines Pfeils) Protokollierung der Begründung ▪ Abfrage der Zusammenhänge zwischen den einzelnen Variablen
Art der Zusammenarbeit	▪ gemeinsame Erarbeitung und Konsentierung in der Gruppe
verwendete Tools	▪ Visualisierung mittels DAGitty ▪ Matrix zur Dokumentation der Zusammenhänge (siehe Anhang E.2.1)
Nachbereitung	▪ Finalisierung des DAG (z. B. Auflösen von Zyklen im DAG) ▪ Ableitung der minimal ausreichenden Adjustierungssets
DAG: gerichteter azyklischer Graph	

Planung und vorbereitende Arbeiten

Es wurde ein 5-stündiger Workshop mit 3 bis 5 praktizierenden Klinikerinnen und Klinikern geplant. Um eine möglichst gute Diskussion unter den Teilnehmenden zu ermöglichen, sollte der Workshop in Präsenz stattfinden. Die Klinikerinnen und Kliniker wurden zum einen über den Ärztlichen Beirat der Deutschen Multiple Sklerose Gesellschaft (DMSG) und zum anderen über die am vorherigen Projekt GA23-02 beteiligte Gruppe externer Sachverständiger akquiriert, schließlich konnten 4 Klinikerinnen und Kliniker für die Teilnahme am Workshop gewonnen werden. Zusätzlich wurde eine Methodikerin mit Expertise in DAGs als externe Sachverständige zur Unterstützung bei der fachlichen Vor- und Nachbereitung, Durchführung und Moderation des Workshops eingebunden. Tabelle 16 listet die eingebundenen Klinikerinnen und Kliniker sowie die Methodikerin inkl. ihrer jeweiligen Affiliation.

Vorab wurde den Klinikerinnen und Klinikern die Fragestellung der vorliegenden Confounderidentifikation, das in GA23-02 identifizierte Confounder-Set und ein Vorschlag zur inhaltlichen Vorsortierung (Clusterung) der Variablen zur Verfügung gestellt.

Tabelle 16: Teilnehmende klinische / methodische Expertinnen und Experten des Workshops

Klinikerin / Kliniker (Affiliation) Friedemann Paul (Charité Berlin) Lisa Ann Gerdes (Institut für Klinische Neuroimmunologie, LMU München) Boris Alexander Kallmann (Kallmann Neurologie & mszb) Tim Woelfle (Research Center for Clinical Neuroimmunology and Neuroscience Basel [RC2NB], University and University Hospital of Basel, Switzerland)
Methodikerin (Affiliation) Jessica L. Rohmann (Quest Center for Responsible Research, Berlin Institute of Health at Charité – Universitätsmedizin Berlin)

Auf eine Abfrage von Interessenkonflikten der eingebundenen Klinikerinnen und Kliniker wurde aufgrund des methodischen Fokus dieses Projekts verzichtet.

Durchführung des DAG-Workshops

Der Workshop wurde entgegen der ursprünglichen Planung in einem hybriden Format durchgeführt, da es einem der Teilnehmenden nicht möglich war vor Ort teilzunehmen.

Zu Beginn erfolgte eine kurze Einführung in die Fragestellung, für welche die Confounderidentifikation durchgeführt wurde. Anschließend folgte eine ausführliche Einführung durch die externe Sachverständige zu den Grundlagen von DAGs und wie Zusammenhänge zwischen Variablen mittels DAG visualisiert werden können. Gleichzeitig wurden verschiedene Begriffe wie Confounder, Collider und die Ableitung eines minimalen Adjustierungssets anhand von Beispielen erläutert. Ziel war, den Teilnehmenden eine Orientierung in der Erarbeitung von DAGs zu geben und vor allem die Hintergründe für deren Anwendung zu erläutern.

Daraufhin wurden den Teilnehmenden die 28 in GA23-02 zusammengeführten Variablen und deren Operationalisierung anhand eines vorbereiteten zeitlichen Rasters (siehe Abbildung 10 in Anhang E.1) vorgestellt, in welchem die Variablen in die Zeiträume: vor MS-Diagnose; nach MS-Diagnose aber vor Therapieentscheidung und zum Zeitpunkt der Therapieentscheidung (Baseline, t_0) gegliedert wurden, um ein gemeinsames Verständnis zu den Variablen herzustellen. Inhaltlich begründete Änderungswünsche an der Operationalisierung oder die Hinzunahme weiterer, aus Kliniker-Sicht wesentlicher Confounder wurden berücksichtigt. Dies wurde als ein zusätzlicher Validierungsschritt des zuvor identifizierten Confounder-Sets angesehen.

Die DAG-Erarbeitung erfolgte durch die getrennte Abfrage des Einflusses der Variablen auf Therapieentscheidung und Endpunkt.

Dabei wurden folgende methodischen Grundsätze definiert, die bei der DAG-Erstellung berücksichtigt werden sollten:

- Wenn für mindestens 1 Patientin / Patient ein kausaler Zusammenhang zwischen 2 Variablen angenommen wurde, wurde ein Pfeil gezeichnet.
- Wenn für eine Variable ein kausaler Zusammenhang auf mindestens 1 Endpunkt angenommen wurde, wurde ein Pfeil gezeichnet.
- Fehlender Konsens oder Unsicherheiten über die Existenz eines Zusammenhangs werden dokumentiert.
- Plausible Gründe für das Weglassen eines Pfeils sind zum Beispiel:
 - Indirekte Wirkung über andere Variablen auf die Therapieentscheidung oder auf den Endpunkt, oder
 - die fehlende Erhebung der Variable zum Zeitpunkt der Therapieentscheidung.
- Das Fehlen von Evidenz ist nicht damit gleichzusetzen, dass kein Zusammenhang besteht. Somit sollte eine einzelne „Nullstudie“ (also eine Studie, die keinen statistisch signifikanten Zusammenhang zwischen den Variablen zeigt) nicht zum Weglassen des Pfeils führen, während mehrere „Nullstudien“ aussagekräftig genug sein können, um einen Pfeil wegzulassen.
- Die Notwendigkeit eines Selektionsknotens ist für jede Fragestellung zu prüfen.

Zum Start der DAG-Erarbeitung wurden die Klinikerinnen und Kliniker für jede Variable gefragt, ob diese ihre Therapieentscheidung beeinflusst. Wenn eine Variable entsprechend der Einschätzung der klinischen Expertinnen und Experten keinen direkten Einfluss auf die Therapieentscheidung hat, wurde geprüft, ob die Variablen diese ggf. indirekt über andere dargestellte oder auch nicht dargestellte Variablen beeinflusst. Wenn dies nicht der Fall war, wurde die Begründung dafür, dass kein Zusammenhang besteht, protokolliert. Die Protokollierung und direkte Visualisierung der besprochenen Zusammenhänge erfolgte mit dem Webtool DAGitty.

Im nächsten Schritt wurde der Einfluss der einzelnen Variablen auf die Endpunkte besprochen. Hier wurde ein pragmatisches Vorgehen gewählt: Alle 4 in der Fragestellung angeführten Endpunkte wurden gemeinsam betrachtet.

Zum Schluss wurden die Zusammenhänge der einzelnen Variablen untereinander besprochen. Hier war der zeitliche Bezug der Variablen zueinander entscheidend. Falls nicht anders

angegeben, war die Ausprägung zu Baseline (t_0 , Tag der Therapieentscheidung) zu berücksichtigen.

Bestand für Zusammenhänge kein Konsens oder wurde der Zusammenhang als unsicher angegeben, wurde dies festgehalten, um mithilfe eines weiteren DAG prüfen zu können, ob die geänderte Pfeilrichtung oder das Weglassen eines Pfeils zu einer Änderung des minimal ausreichenden Adjustierungssets führt. Falls durch einen solchen weiteren DAG ein anderes minimales Adjustierungsset entsteht, sollte dieses für eine Sensitivitätsanalyse genutzt werden. Um weggelassene Pfeile von noch nicht besprochenen Zusammenhängen abzugrenzen, wurden diese gesondert dokumentiert und der Grund für das Weglassen eines Pfeils zur Therapieentscheidung oder Endpunkt festgehalten.

Da der DAG nicht in der angesetzten Workshop-Zeit finalisiert werden konnte, wurde ein weiterer virtueller Termin zur Finalisierung des DAG angesetzt. An der Fortsetzung nahmen 2 der 4 Klinikern und Kliniker sowie die externe Sachverständige teil. Der virtuelle Workshop dauerte 2,5 Stunden. Insgesamt summierte sich die benötigte Zeit für die Durchführung des Workshops somit auf 7,5 Stunden. Nicht berücksichtigt ist dabei die Zeit der internen Vor- und Nachbereitung.

Zur Vorbereitung der Workshop-Fortsetzung wurden durch die IQWiG-Projektgruppe bereits Annahmen für einige der noch offenen Zusammenhänge getroffen, welche den Klinikern im Vorfeld zur Überprüfung zugeschickt und im Falle getroffener, nicht plausibler Annahmen mit den Klinikern konsentiert wurden. Um sicherzustellen, dass alle Zusammenhänge systematisch besprochen wurden, erfolgte die Dokumentation anhand einer Matrix (siehe Anhang E.2.1, Abbildung 11). Im Nachgang wurden die Zusammenhänge mittels DAGitty visualisiert und minimal ausreichende Adjustierungssets abgeleitet.

5.2.1.7.3 Ergebnis des DAG-Workshops

Inhaltliche Ergebnisse

Das in GA23-02 identifizierte Confounder-Set umfasste insgesamt 28 Variablen. Bei den im Workshop eingebrachten begründeten Änderungen der Klinikern und Kliniker handelte es sich nicht um die Aufnahme neuer Confounder, sondern um die separate Betrachtung einzelner Teilkomponenten der zuvor zusammengeführten Confounder, wie z. B. Zeit seit der Erstmanifestation oder die Berücksichtigung der Veränderung der Krankheitsaktivität im Vorjahr. Insgesamt führten die Änderungen zu einem Confounder-Set aus 31 Variablen (siehe Tabelle 27 in Anhang E.4).

Die Begründungen für weggelassene Pfeile zur Therapieentscheidung und Endpunkten sind in Tabelle 17 dokumentiert. Abweichend zur Einschätzung der Klinikern und Klinikern

wurden im Nachgang zum Workshop die weggelassenen Pfeile von CSF Status und Patientenwunsch zum Endpunkt dennoch eingezeichnet. Begründet ist dies darin, dass für den CSF Status die Evidenzlage nicht ausreichend eindeutig ist, um den Ausschluss des Pfeils zu rechtfertigen [71-74]. Für die Variable Patientenwunsch ist davon auszugehen, dass der Wunsch einer Patientin / eines Patienten sich zumindest auf den Endpunkt der gesundheitsbezogenen Lebensqualität auswirkt.

Tabelle 17: Dokumentation der weggelassenen Pfeile zur Therapieentscheidung und zu den Endpunkten

	Confounder	Begründung
Kein direkter Einfluss (Pfeil) auf die Therapieentscheidung	Abstammung	Indirekter Einfluss über die Variablen Bildungsgrad und Zeit seit der Erstmanifestation auf die Therapieentscheidung.
	CSF Status	Erhebung erfolgt nur zur Diagnose (siehe MS-Leitlinie [75]) und nicht zur Therapieentscheidung.
	Tabakkonsum	Indirekter Einfluss über die Variablen Komorbiditäten, Supernode Veränderung der Krankheitsaktivität im Vorjahr und den Supernode Krankheitsaktivität Baseline.
	Fatigue	Kein direkter Einfluss auf Therapieentscheidung, da keins der beiden Arzneimittel die Fatigue verstärkt.
Kein direkter Einfluss (Pfeil) auf die Endpunkte	Art der beruflichen Tätigkeit	Variable wirkt indirekt über andere Variablen wie gesundheitsbezogene Lebensqualität auf die Endpunkte.
	Bildungsgrad	Variable wirkt indirekt über andere Variablen wie Tabakkonsum und BMI auf die Endpunkte.
	Verträglichkeit der Vortherapie	Variable wirkt indirekt über absolute Lymphozytenzahl und den Supernode Krankheitsaktivität auf die Endpunkte.
	vorherige Behandlung mit einer anderen DMT	Variable wirkt indirekt über den Supernode Krankheitsaktivität im Vorjahr auf die Endpunkte.
CSF: Cerebrospinalflüssigkeit; DMT: krankheitsmodifizierende Therapie		

Im Prozess der DAG Erstellung wurden einige Confounder zu sogenannten Supernodes gruppiert, was die Übersichtlichkeit im DAG verbesserte. Der Supernode Krankheitsaktivität Baseline umfasst 4 und der Supernode Veränderung Krankheitsaktivität im Vorjahr 7 Confounder. Allen in den Supernodes enthaltenen Confoundern wurde ein vergleichbarer Stellenwert beigemessen, d. h. es gibt keine Variablen, die nachgelagert zueinander sind und alle in den Supernodes enthaltenen Variablen müssen erhoben werden.

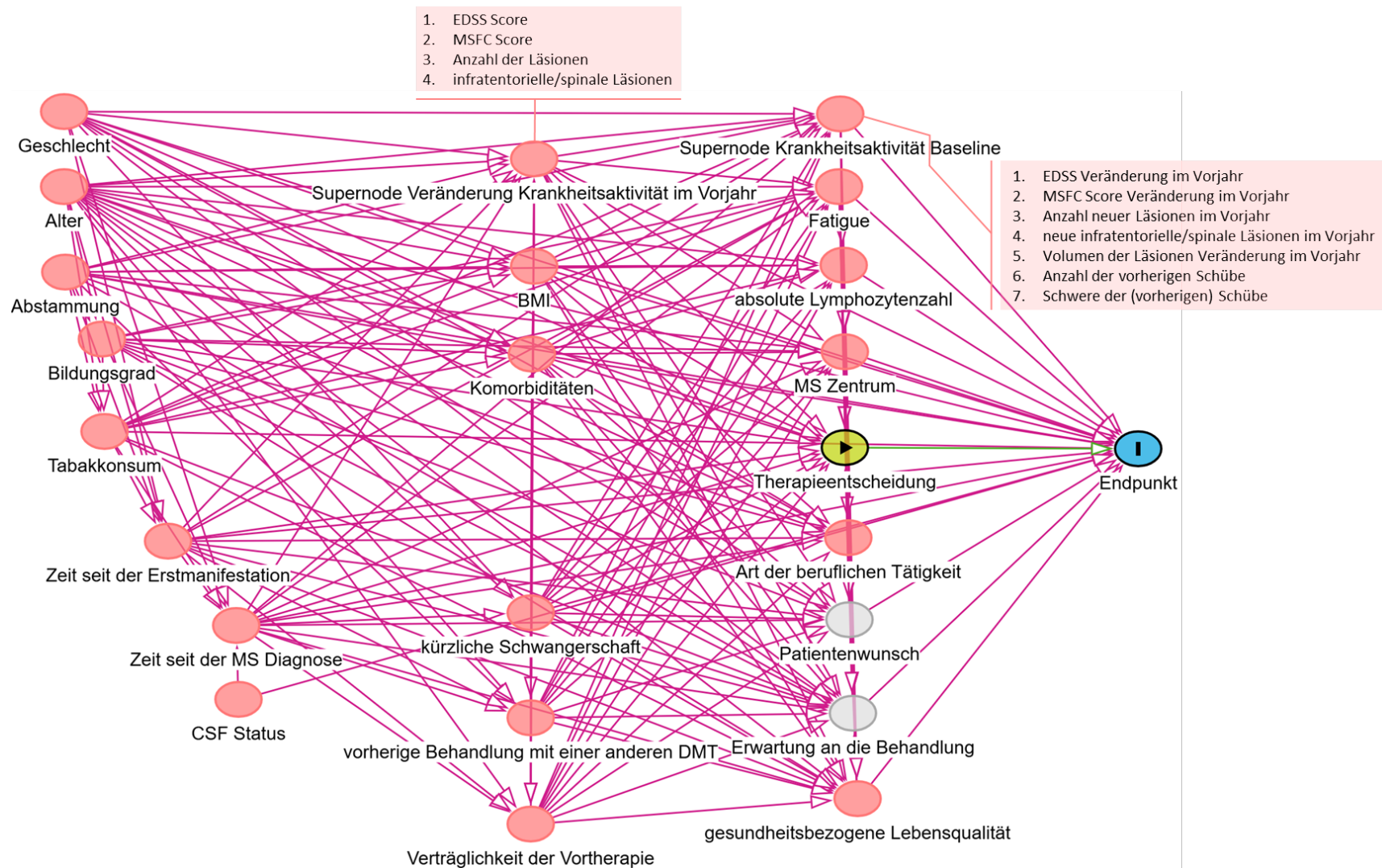
Zusätzlich wurde für die vorliegende Fragestellung festgehalten, dass nicht von konkurrierenden Ereignissen auszugehen ist, welche im DAG berücksichtigt werden müssten. Auch auf die Darstellung des Selektionsknotens „MS-Diagnose“ im DAG wurde verzichtet, da die Patientinnen und Patienten mit MS die Grundgesamtheit bilden und somit nicht von einer

Selektion auszugehen ist. Falls die Fragestellung beispielsweise mit Registerdaten beantwortet werden soll, ist zu prüfen ob die Patientenpopulation im Register repräsentativ für die zu untersuchende Patientenpopulation der Fragestellung ist.

Bei der Besprechung der Zusammenhänge der einzelnen Variablen untereinander wurden insgesamt 4 Zusammenhänge von den Klinikerinnen und Klinikern als unsicher angegeben und in der Matrix entsprechend gekennzeichnet (siehe Abbildung 11). Der Ein- oder Ausschluss dieser 4 als unsicher angegebenen Pfeile führte allerdings zu keiner Veränderung der minimal ausreichenden Adjustierungssets. Aus diesem Grund wird trotz der unsicheren Zusammenhänge nur 1 finaler DAG gezeichnet (unter Annahme des Vorhandenseins der jeweiligen Zusammenhänge).

Da die Visualisierung und Analyse der besprochenen Zusammenhänge im Nachgang der Workshop-Fortsetzung erfolgten, war erst dann mittels DAGitty ersichtlich, dass Zyklen vorlagen (also Pfade, die ausgehend von 1 Variable über andere Variablen wieder zur Ausgangsvariable zurückführen, was entsprechend der DAG-Theorie nicht erlaubt ist). Diese wurden ohne erneute Einbindung der Klinikerinnen und Kliniker in der Projektgruppe und unter Berücksichtigung des zeitlichen Aspekts (auf welchen Zeitpunkt die Variablen sich beziehen) aufgelöst. Die nachträglichen Änderungen sind in Anhang E.5, Tabelle 28 dokumentiert.

Abbildung 7 zeigt den finalen, mittels DAGitty erstellten DAG inklusive der in den Supernodes enthaltenen Confounder. Die beiden Confounder Erwartung an die Behandlung und Patientenwunsch sind grau dargestellt, weil diese bereits im Arbeitspapier GA23-02 als nicht mess- bzw. operationalisierbar bewertet wurden.



BMI: Body-Mass-Index; CSF: Cerebrospinalflüssigkeit; DMT: krankheitsmodifizierende Therapie; EDSS: Expanded Disability Status Scale; MS: multiple Sklerose; MSFC: Multiple Sclerosis Functional Composite

Abbildung 7: Finaler gerichteter azyklischer Graph (DAG)

Mittels des DAG wurden 3 minimal ausreichende Adjustierungssets identifiziert, welche jeweils 27 Confounder umfassen und nachfolgend in Tabelle 18 gelistet sind. Zu berücksichtigen ist dabei, dass für einige Variablen (wie z. B. den EDSS-Score) sowohl die Veränderung im Vorjahr als auch der Wert zu Baseline einen relevanten Confounder darstellen. Da beide Operationalisierungen über das gleiche Datenfeld erhoben werden können, hat dies zur Folge, dass in der vorliegenden Fragestellung für 8 Confounder lediglich 4 Datenfelder erforderlich sind und die 27 Confounder mithilfe von 23 Datenfeldern erfasst werden können.

In jedem Set konnten somit 4 der ursprünglich 31 Variablen eingespart werden. Jedes Confounder-Set für sich wäre theoretisch für eine vollständige Adjustierung ausreichend. Da jedoch in allen Sets die beiden nicht messbaren Confounder Erwartung an die Behandlung und Patientenwunsch enthalten sind, ist auch nach der Adjustierung für alle erhebbaren Confounder davon auszugehen, dass residuelles Confounding vorliegen wird.

Tabelle 18: Darstellung der 3 identifizierten minimal ausreichenden Adjustierungssets^a

Confounder Set-Nr. 1	Confounder Set-Nr. 2	Confounder Set-Nr. 3
▪ Geschlecht ▪ Alter ▪ Fatigue ▪ Komorbiditäten ▪ absolute Lymphozytenzahl ▪ MS Zentrum ▪ BMI ▪ Tabakkonsum ▪ Erwartung an die Behandlung ▪ Patientenwunsch ▪ gesundheitsbezogene Lebensqualität ▪ Zeit seit der MS Diagnose ▪ Bildungsgrad ▪ Abstammung ▪ kürzliche Schwangerschaft ▪ Zeit seit der Erstmanifestation ▪ Supernode Krankheitsaktivität Baseline^b ▪ Supernode Veränderung der Krankheitsaktivität im Vorjahr^c	▪ Geschlecht ▪ Alter ▪ Fatigue ▪ Komorbiditäten ▪ absolute Lymphozytenzahl ▪ MS Zentrum ▪ BMI ▪ Tabakkonsum ▪ Erwartung an die Behandlung ▪ Patientenwunsch ▪ gesundheitsbezogene Lebensqualität ▪ Zeit seit der MS Diagnose ▪ CSF Status ▪ Abstammung ▪ kürzliche Schwangerschaft ▪ Zeit seit der Erstmanifestation ▪ Supernode Krankheitsaktivität Baseline^b ▪ Supernode Veränderung der Krankheitsaktivität im Vorjahr^c	▪ Geschlecht ▪ Alter ▪ vorherige Behandlung mit einer anderen DMT ▪ Komorbiditäten ▪ absolute Lymphozytenzahl ▪ MS Zentrum ▪ BMI ▪ Verträglichkeit der Vorthherapie ▪ Erwartung an die Behandlung ▪ Patientenwunsch ▪ gesundheitsbezogene Lebensqualität ▪ Zeit seit der MS Diagnose ▪ Bildungsgrad ▪ Art der beruflichen Tätigkeit ▪ kürzliche Schwangerschaft ▪ Zeit seit der Erstmanifestation ▪ Supernode Krankheitsaktivität Baseline^b ▪ Supernode Veränderung der Krankheitsaktivität im Vorjahr^c
a. Variablen, die sich zwischen den Confounder-Sets unterscheiden, sind fett markiert. b. Der Supernode Krankheitsaktivität Baseline umfasst folgende 4 Confounder: EDSS Score, MSFC Score, Anzahl der Läsionen, infratentorielle / spinale Läsionen. c. Der Supernode Veränderung der Krankheitsaktivität im Vorjahr umfasst folgende 7 Confounder: EDSS Veränderung im Vorjahr, MSFC Score Veränderung im Vorjahr, Anzahl neuer Läsionen im Vorjahr, neue infratentorielle / spinale Läsionen im Vorjahr, Volumen der Läsionen Veränderung im Vorjahr, Anzahl der vorherigen Schübe, Schwere der (vorherigen) Schübe. BMI: Body-Mass-Index; CSF: Cerebrospinalflüssigkeit; DMT: krankheitsmodifizierende Therapie; EDSS: Expanded Disability Status Scale; MS: multiple Sklerose; MSFC: Multiple Sclerosis Functional Composite		

Methodische Eignung eines DAG-Workshops für die systematische Confounderidentifikation

In der Literatur existieren nur wenige Publikationen dazu, wie die praktische Einbindung von Experten(wissen) in die DAG-Erstellung erfolgen soll. Grundsätzlich gibt es verschiedene Möglichkeiten, diese Einbindung auszugestalten. Der für dieses Projekt geplante Ablauf (wie z. B. der Start mit einem leeren [kantenlosen] anstelle eines gesättigten DAG) stellen einen von mehreren möglichen Wegen dar. Als Grundlage für die Ableitung von konkreten Handlungsempfehlungen (siehe Abschnitt 5.3.3) werden im Folgenden die sich in der

Durchführung des Workshops als hilfreich oder auch weniger hilfreich gezeigten Aspekte beschrieben und diskutiert:

- Der in Präsenz durchgeführte Workshop stellte sich als eine sehr diskussions- und konsens-fördernde Lösung zur aktiven Einbindung der klinischen Expertinnen und Experten heraus.
- Für den Workshop wurden 4 klinische Expertinnen und Experten eingebunden, was innerhalb des geplanten Rahmens von 3-5 Klinikerinnen und Klinikern lag. Diese Anzahl war angemessen, um ein breites Spektrum an Expertise abzudecken.
- Besonders hilfreich für die Erarbeitung des DAG war die Einbindung einer Methodikerin mit Expertise und Erfahrung in der Erstellung von DAGs, insbesondere in Hinblick auf die Moderation der Diskussion.
- Ein zeitlicher Rahmen von 1 Tag war für die Erstellung eines DAG von 28 Variablen angemessen, aber auch notwendig. Die im vorliegenden Fall notwendige Aufteilung auf 2 halbe Tage hatte unter anderem den Vorteil, methodische Aspekte oder ggf. Annahmen zu rekapitulieren, spezifische Literatur zu recherchieren oder auch weitere Unterlagen vorzubereiten.
- Vor dem Start der eigentlichen DAG-Erarbeitung ist es wesentlich, eine Einführung in die Theorie von DAGs zu geben. Dies hat im vorliegenden Workshop ca. 45 Minuten in Anspruch genommen. Dieses gemeinsame Basis-Verständnis der Grundsätze und Zielstellung der DAG-Erstellung hat sich im Verlauf der Diskussion immer wieder als zielführend bestätigt.
- Bevor mit der DAG-Erarbeitung gestartet wird, ist es wesentlich, den Teilnehmerinnen und Teilnehmern bereits im Vorfeld des Workshops die zugrunde liegende Fragestellung und die vorgeschlagene Operationalisierung der einzelnen Variablen zur Verfügung zu stellen und zu Beginn des Workshops mit den eingebundenen Klinikerinnen und Klinikern zu besprechen und damit ein gemeinsames Verständnis der Variablen zu erlangen. Dieser Schritt stellt gleichzeitig eine Validierung dar – dass keine relevante Variable im Confounder-Set fehlt und auch dass alle Variablen im Versorgungsalltag (sofern möglich) in der vorgeschlagenen Operationalisierung erhoben werden.
- Bei der Operationalisierung der Variablen ist insbesondere der zeitliche Aspekt, d. h. auf welchen Erhebungszeitpunkt die Variablen sich beziehen, zu berücksichtigen. Ein zeitliches Raster kann hier bei der Visualisierung hilfreich sein (siehe Anhang E.1). Das zeitliche Raster wurde in dem vorliegenden Fall einer Clusterung der Variablen in 3 inhaltlichen Gruppen vorgezogen, da dieses im Vergleich zur Erstellung eines DAG pro Cluster und der Kombination dieser Cluster als zeitsparender eingeschätzt wurde.

- Der Start mit einem leeren (kantenlosen) DAG war sehr zeitaufwendig. Ggf. kann der Ansatz, mit einem gesättigten DAG (in dem alle möglichen Pfeile eingezeichnet sind und anschließend in der Diskussion wieder entfernt werden) Zeit sparen. In jedem Fall war die Dokumentation der besprochenen Zusammenhänge in einer Matrix sehr hilfreich, um sicherzustellen, dass alle Zusammenhänge besprochen wurden, und ermöglichte gleichzeitig ein systematisches und strukturiertes Vorgehen (siehe Anhang E.2.1).
- Ziel war, einen Konsens für die Zusammenhänge zwischen den Variablen zu erzielen. In einzelnen Fällen war dies nicht möglich. In der Diskussion war es hilfreich, den Klinkerinnen und Klinkern zu verdeutlichen, dass selbst bei fehlendem Konsens zu einzelnen Zusammenhängen (und damit der Zeichnung eines 2. DAG) die Konsequenz nicht unbedingt ein unterschiedliches minimal ausreichendes Adjustierungsset sein muss. Je nach Konstellation kann der Dissens konsequenzlos sein.
- An verschiedenen Stellen wurde bei der DAG-Erarbeitung deutlich, dass ein DAG grundsätzlich auf einer Reihe getroffener Annahmen basiert, welche je nach Gruppe eingebundener Klinikerinnen und Kliniker auch voneinander abweichen können. Umso wichtiger ist eine transparente Dokumentation mindestens der Gründe für weggelassene Pfeile zur Therapieentscheidung oder Endpunkten (da – wie oben beschrieben – das Weglassen eines Pfeils die deutlich stärkere Annahme darstellt als das Zeichnen eines Pfeils).

Die konkreten Handlungsempfehlungen, die sich daraus ergeben, sind in Abschnitt 5.3 zusammengefasst.

5.2.1.7.4 Zusammenfassung

Insgesamt wurde in dem Workshop unter Experteneinbindung ein DAG bestehend aus 31 Variablen für die vorliegende Fragestellung erstellt. Hierbei lag der Fokus insbesondere auf der methodischen Vorgehensweise (Proof-of-Concept) hinsichtlich der DAG Erstellung und der notwendigen transparenten Dokumentation dieses Vorgehens. Durch die Erarbeitung des DAG wurden im vorliegenden Beispiel 3 minimal ausreichende Adjustierungssets identifiziert und 4 der 31 Confounder konnten eingespart werden. Zur Erhebung der 27 Confounder werden 23 Datenfelder in einer Datenquelle benötigt. Der damit verbundene Aufwand zur Erarbeitung des DAG (Dauer des Workshops insgesamt 7,5 Stunden) wird als überschaubar eingeschätzt. Im vorliegenden Beispiel war mit 4 Confoundern zwar lediglich eine geringe Aufwandsreduktion möglich, allerdings besteht das Potenzial je nach Fragestellung bzw. Indikation eine höhere Anzahl an Confoundern einzusparen. Ein wesentlicher Vorteil der Erstellung eines DAG ist die transparente Dokumentation der zugrunde liegenden Annahmen. Unsicherheiten bezüglich getroffener Annahmen können gekennzeichnet und deren potenzielle Auswirkung auf das minimal ausreichende Adjustierungsset überprüft werden.

Aus den gewonnenen Erfahrungen werden übergeordnete Handlungsempfehlungen abgeleitet, welche in Abschnitt 5.3 beschrieben werden.

5.2.1.8 Relevanz der Confounder – Bewertung der Stärke des Einflusses auf die Effektschätzung

Das Ziel der im vorhergehenden Kapitel 5.2.1.7 beschriebenen DAG-Erstellung ist, diejenige Teilmenge der identifizierten Confounder zu ermitteln, für die mindestens adjustiert werden muss, um alle identifizierten Confounder adäquat in der Auswertung zu berücksichtigen. Die in diesem Schritt in den minimal ausreichenden Adjustierungssets identifizierten Confounder können jedoch einen deutlich unterschiedlichen Einfluss auf die Effektschätzung haben. Damit ist ein weiterer inhaltlich begründeter Ansatz ein Confounder-Set zu reduzieren, nur diejenigen Confounder auszuwählen, die den größtmöglichen Einfluss auf die Effektschätzung haben bzw. diejenigen aus dem Set zu entfernen, die den geringsten Einfluss aufweisen. Ein möglicher Ansatz sich dem Einfluss der Confounder auf die Effektschätzung anzunähern, ist, die inhaltliche Einschätzung der Relevanz der identifizierten Confounder zu ermitteln.

Der Ansatz, identifizierte Confounder nach ihrer Relevanz zu bewerten, ist bereits an verschiedenen Stellen beschrieben worden. So haben Pufulete 2022 klinische Expertinnen und Experten gebeten, aus der identifizierten Confounder-Liste die 5 Wichtigsten zu benennen. In GA23-02 wurde dieser Vorschlag dahin gehend erweitert, als dass im Sinne eines höheren Informationsgehalts eine Punktevergabe für jeden einzelnen Confounder empfohlen wurde.

Bisher ist dieser Schritt der Relevanzbewertung nicht zur Reduktion eines Confounder-Sets empfohlen worden. Lediglich in den 3 zuvor durch die Literaturrecherche identifizierten Publikationen [32-34], die zumindest in Teilen das Thema der Reduktion eines Confounder-Sets behandeln (siehe Abschnitt 5.2.1), wurde die Relevanzbewertung als mögliche Entscheidungshilfe für die Confounder-Selektion diskutiert.

Um zu prüfen, wie eine Relevanzbewertung zur Reduktion von Confoundern beitragen kann, wurde im vorliegenden Projekt – analog zum Vorgehen der externen Sachverständigen in GA23-02 und den 3 identifizierten Publikationen – eine Umfrage zur Bewertung der Stärke des Einflusses der identifizierten Confounder durchgeführt. Die Bewertung erfolgte durch klinische Expertinnen und Experten mit Expertise in der entsprechenden Indikation (hier RRMS). Die Methodik und die Ergebnisse der Umfrage werden im Folgenden beschrieben.

Ablauf und Aufbau der Umfrage

Ausgangspunkt für die Umfrage bildete das Ergebnis des in Abschnitt 5.2.1.7.3 beschriebenen DAG-Workshops. Da die 3 identifizierten minimal ausreichenden Adjustierungssets zusammen alle Confounder umfassen, wurden alle 31 Confounder für die Bewertung eingeschlossen. Ein

Confounder, der in keinem der Adjustierungssets enthalten ist, könnte aus der Umfrage ausgeschlossen werden.

Um den Teilnehmenden die Bewertung zu vereinfachen, wurden die 31 Confounder vorab in verschiedene thematische Gruppen (z. B. demographisch, sozioökonomisch, Lebensstil und Komorbiditäten [siehe auch Umfrage in Anhang F.2]) unterteilt. Für jeden Confounder wurde eine eindeutige Operationalisierung angegeben. Zusätzlich wurde in der Umfrage darauf hingewiesen, dass – sofern nicht anders angegeben – die Ausprägung des jeweiligen Confounders zum Zeitpunkt der Therapieentscheidung zu berücksichtigen ist und jeder Confounder im Kontext des kompletten Confounder-Sets zu bewerten ist. Für jeden Confounder wurde auf einer 5-Punkte-Skala von 1 („sehr gering“) bis 5 („sehr stark“) sowohl die Stärke des Einflusses auf die Therapieentscheidung als auch auf die Endpunkte abgefragt. Dabei wurde der Einfluss auf die Endpunkte Schubrate, Behinderungsprogression, gesundheitsbezogene Lebensqualität und unerwünschte Ereignisse gemeinsam betrachtet. Falls die Einschätzung für die einzelnen Endpunkte unterschiedlich ausfiel, sollte die Einschätzung für den Endpunkt angegeben werden, der am stärksten beeinflusst wird. In der Umfrage wurde darum gebeten, wann immer möglich, die gesamte Spanne der 5 Kategorien auszuschöpfen, um eine Differenzierung hinsichtlich der Relevanz zu ermöglichen. Die Umfrage ist in Anhang F.2 abgebildet.

Um eine möglichst hohe Rücklaufquote zu erzielen, wurde die Umfrage online durchgeführt. Den Teilnehmenden wurde in einem Begleitschreiben (siehe Anhang F.1) eine kurze Beschreibung des Projekts und der vorliegenden Fragestellung zur Verfügung gestellt.

Empfängerinnen und Empfänger waren der Ärztlichen Beirat der DMSG, und zusätzlich die Ärztlichen Beiräte und MS-Zentren der DMSG-Landesverbände. Die Laufzeit der Umfrage betrug 5 Wochen. Die Umfrage erfolgte anonym.

Während der laufenden Umfrage wurden von einem der Teilnehmenden Unklarheiten zurückgemeldet, da die Fragestellung (Dimethylfumarat vs. Glatirameracetat) nicht mehr den aktuellen Versorgungskontext abbildete. Um zu vermeiden, dass dies trotz Erläuterung im Begleitschreiben (siehe Anhang F.1) zu Unklarheiten bei weiteren Teilnehmenden führt, wurde ca. 3 Wochen nach Start der Umfrage eine erläuternde E-Mail an die Klinikerinnen und Kliniker verschickt. In dieser wurde verdeutlicht, dass die Klinikerinnen und Kliniker sich für die Umfrage in die Situation zurückversetzen sollten, als sie für eine Ihrer erwachsenen Patientinnen bzw. einen Ihrer erwachsenen Patienten mit RRMS die Entscheidung treffen mussten, ob sie Dimethylfumarat oder Glatirameracetat verordnen unter der Voraussetzung, dass die neueren Therapieoptionen zu diesem Zeitpunkt noch nicht zur Verfügung standen. Alle, die die Umfrage bereits mit einem anderen Verständnis der Fragestellung ausgefüllt hatten, wurden gebeten, dies erneut zu tun. An der Umfrage selbst wurden keine Änderungen vorgenommen.

Ergebnis der Online-Umfrage

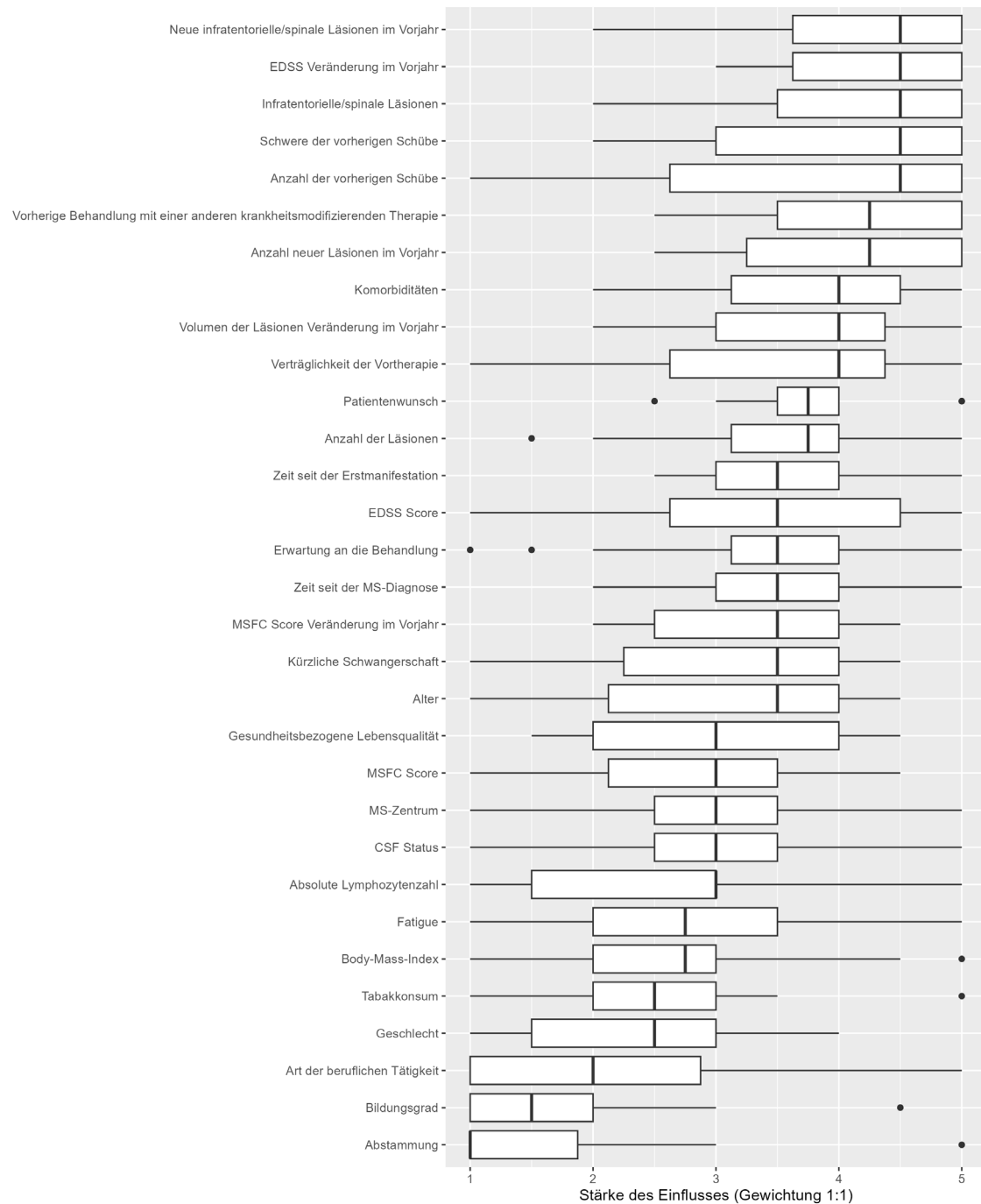
Insgesamt nahmen 18 Klinikerinnen und Klinikern an der Umfrage teil. Alle Teilnehmenden füllten die Umfrage vollständig aus.

Um ein Ranking der Confounder zu ermöglichen, ist es notwendig die beiden getrennten Bewertungen (Stärke des Einflusses auf die Therapieentscheidung bzw. Endpunkte) zusammenzufassen. Hierzu gibt es verschiedene Möglichkeiten. Zum einen können beide Bewertungen gleichwertig mit einer Gewichtung von 1:1 gemittelt werden. Zum anderen sind auch weitere Gewichtungen von z. B. 2:3 und 1:2 zugunsten der Stärke des Einflusses auf den Endpunkt vorstellbar, da verschiedene Arbeiten im Kontext der Relevanz von Confoundern auf die vorrangige Bedeutung des Einflusses eines Confounders auf den Endpunkt hinweisen [51,76,77].

Dementsprechend wurden bei der Auswertung der vorliegenden Umfrage die Bewertungen je Confounder sowohl mit einer Gewichtung von 1:1 gemittelt als auch die weiteren Gewichtungen 2:3 und 1:2 getestet. Das Ranking der Confounder wurde anhand des Medians vorgenommen. Da das Ranking auch unter Anwendung von Gewichtungen zugunsten der Stärke des Einflusses auf den Endpunkt nur marginal beeinflusst wurde, wurden diese für die weiteren Betrachtungen nicht mehr berücksichtigt. Im Sinne eines pragmatischen Vorgehens wird im Folgenden nur die 1:1 Gewichtung beschrieben (die Ergebnisse der weiteren Gewichtungsverhältnisse sind in Tabelle 29 im Anhang F.3 dargestellt).

Abbildung 8 veranschaulicht die Ergebnisse des Rankings mit der Gewichtung 1:1 in Form von Boxplots für jeden einzelnen Confounder. Die Confounder sind darin entsprechend der medianen Stärke ihres Einflusses von „sehr stark“ (5 Punkte) bis „sehr gering“ (1 Punkt) sortiert. Die vollständigen Ergebnisse der Umfrage sind in Tabelle 29 im Anhang F.3 dargestellt.

Lediglich 4 Klinikerinnen und Kliniker füllten die Umfrage nach der per E-Mail versendeten Klarstellung aus. Auf Sensitivitätsanalysen zur Untersuchung möglicher Auswirkungen der Klarstellung auf das Umfrageergebnis wurde daher verzichtet. Dies wird vor dem Hintergrund des primären Ziels der Durchführbarkeit des Vorgehens zur Relevanzbewertung als vertretbar angesehen.



Der Rechteckbereich zeigt die Werte vom 1. bis zum 3. Quartil und die darin enthaltene horizontale Linie den Median an. Die horizontalen Linien (Whisker) erstrecken sich vom jeweiligen Quartil (Q1, Q3) bis zu demjenigen Wert, der höchstens 1,5 x IQR entfernt ist. Werte darüber hinaus werden als Punkte dargestellt. BMI: Body-Mass-Index; CSF: Cerebrospinalflüssigkeit; DMT: krankheitsmodifizierende Therapie; EDSS: Expanded Disability Status Scale; IQR: Interquartilsabstand; MS: multiple Sklerose; MSFC: Multiple Sclerosis Functional Composite; Q1: 1. Quartil; Q3: 3. Quartil

Abbildung 8: Boxplots der Ergebnisse der Online-Umfrage (n = 18). Ranking der Confounder entsprechend ihrer Einflussstärke von „sehr stark“ (5) bis „sehr gering“ (1) bei einer Gewichtung der Stärke des Einflusses auf Therapieentscheidung bzw. auf die Endpunkte von 1:1

Insgesamt wurde die Stärke des Einflusses von 10 der 31 Variablen durch die Klinikerinnen und Kliniker im Median mit ≥ 4 Punkten (d. h. stark oder sehr stark) bewertet. Vielen Confoundern der Kategorie Krankheitsschwere wurde durch die Klinikerinnen und Kliniker eine hohe Bedeutung beigemessen (7 der 10 am höchsten Bewerteten). Die Variablen Art der beruflichen Tätigkeit, Bildungsgrad und Abstammung wurden im Median mit ≤ 2 Punkten (d. h. gering oder sehr gering) bewertet.

Anwendung des Rankings zur Reduktion der Confounder

Sollte eine Reduktion des über die vorherigen Schritte identifizierten Confounder-Sets notwendig sein, kann das aus der Umfrage abgeleitete Ranking der Confounder als Grundlage dienen. Hier sind verschiedene Ansätze vorstellbar, wie das Set an Confoundern reduziert werden kann. Die zugrunde liegende Annahme ist, dass die Höhe des Punktwertes mit der Größe des Einflusses auf die Effektschätzung korreliert. Eine Möglichkeit ist, einen Cut-off für den (medianen) Punktwert festzulegen. Denjenigen Confoundern, deren medianer Punktwert unter diesem Cut-off liegt, würde dann nur eine untergeordnete Relevanz beigemessen, sodass der Ausschluss dieser Confounder aus dem finalen Set vertretbar ist. Ein anderer Ansatz ist, vorab eine konkrete Zahl an Confoundern zu definieren, die aus der Liste entfernt werden soll (z. B. das Fünftel der Liste mit den niedrigsten Punktwerten).

Aus inhaltlichen Überlegungen heraus ist es sinnvoll, einen Cut-off zu wählen, der sich an der erwarteten Bedeutung der Confounder orientiert. Dadurch werden diejenigen Confounder aus dem finalen Set entfernt, denen der geringste Einfluss beigemessen wird. Eine vorab definierte Anzahl an Confoundern aus dem Confounder-Set zu entfernen, kann hingegen dazu führen, dass auch Confounder ausgeschlossen werden, denen eine hohe Bedeutung beigemessen wird. Dies wäre nicht sachgerecht.

Durch das Anlegen eines Cut-offs von ≤ 2 Punkten im Median werden diejenigen Confounder aus dem finalen Set entfernt, deren Stärke des Einflusses als sehr gering (1 Punkt) oder gering (2 Punkte) bewertet wurde und denen somit der geringste Einfluss beigemessen wird. Dies führt im vorliegenden Confounder-Set zu einer Reduktion von 3 Confoundern (Abstammung, Bildungsgrad und Art der beruflichen Tätigkeit). Das Entfernen dieser 3 Confounder aus den minimal ausreichenden Adjustierungssets hat zur Folge, dass statt der ursprünglich 3 identifizierten minimal ausreichenden Adjustierungssets mit je 27 Confoundern (siehe Tabelle 18) nur noch 2 Sets (Set 1 und Set 3) mit je 25 Confoundern bestehen (für die je 21 Datenfelder in einer Datenquelle benötigt werden).

5.2.2 Zusammenfassende Gegenüberstellung geeigneter und nicht geeigneter Maßnahmen zur Reduktion eines Confounder-Sets

Das Ziel der Reduktion eines Confounder-Sets muss immer im Kontext mit dem damit einhergehenden Verlust an Ergebnissicherheit betrachtet werden. Ein Ausschluss von

Confoundern sollte deshalb immer so weit es geht inhaltlich begründet sein. Die in den Abschnitten 5.2.1.1 bis 5.2.1.6 beschriebenen Maßnahmen (Schwellenwerte, Berücksichtigung der Art oder Qualität der Quelle, Korrelationen, Fokussierung auf Faktoren die die Therapieentscheidung beeinflussen, Verzicht auf instrumentelle Variablen) haben sich als nicht zielführend für die Reduktion des Confounder-Sets erwiesen, da diese nicht ausreichend inhaltlich begründet sind.

Einzig die Erarbeitung eines DAG (Abschnitt 5.2.1.7) und ggf. die anschließende Bewertung der Confounder hinsichtlich der Stärke des Einflusses auf die Therapieentscheidung und die Endpunkte (und damit der vermuteten Stärke des Einflusses auf die Effektschätzung) (Abschnitt 5.2.1.8) stellen geeignete Maßnahmen zur systematischen und transparenten Reduktion eines identifizierten Confounder-Sets dar. In einem DAG werden die der Fragestellung zugrunde liegenden Annahmen über die qualitativen kausalen Zusammenhänge explizit und damit transparent gemacht und die Reduktion basiert auf inhaltlichen Überlegungen. Auch wenn das DAG im vorliegenden Beispiel nicht zu einer relevanten Reduktion an zu erhebenden Confoundern geführt hat, ist es mindestens ein wichtiger Schritt zur Validierung des identifizierten Confounder-Sets und wird entsprechend als Maßnahme empfohlen. In anderen Fragestellungen kann mithilfe eines DAG ggf. eine relevante Reduktion an Confoundern erzielt werden.

Bei der DAG-Erstellung und der Bestimmung minimaler Adjustierungssets wird entsprechend der Methodik nicht berücksichtigt, wie stark ein Confounder auf die Therapieentscheidung bzw. den Endpunkt wirkt. Die Berücksichtigung dieser Stärke des Einflusses kann im letzten Schritt eine weitere Maßnahme zur Reduktion des Confounder-Sets darstellen. So können auf Basis der Einschätzung von klinischen Expertinnen und Experten die Confounder aus dem finalen Confounder-Set entfernt werden, denen nur ein untergeordneter Einfluss beigemessen wird. Entscheidend ist hier die Wahl des Cut-offs, der zur Reduktion des Confounder-Sets genutzt wird. Dabei ist es vertretbar, Confounder, denen nur eine sehr geringe und geringe Bedeutung beigemessen wird, aus dem finalen Confounder-Set zu entfernen. Wird die Mehrzahl an Confoundern als wichtig eingeordnet, ist keine relevante Aufwandsreduktion möglich, ohne das Risiko eines nicht abschätzbaren Verlusts an Ergebnissicherheit einzugehen. Im vorliegenden Beispiel hat die Anwendung dieses letzten Schritts ebenfalls nur zu einer sehr geringen Aufwandsreduktion hinsichtlich der Datenerhebung geführt.

Aufgrund des potenziell nur geringen Einsparpotenzials und der zusätzlichen Unsicherheiten, die dieser Schritt mit sich bringt, sollte dieser Schritt nur in Abhängigkeit der Anzahl an Confoundern und der für die Studie zur Verfügung stehenden Patientenzahl erfolgen. Der Schritt der Reduktion (siehe Schritt 3b im nachfolgenden Abschnitt) ist in Situationen nicht notwendig, in denen eine ausreichende Anzahl an Patientinnen und Patienten für die Studie

zur Verfügung steht und die Anzahl identifizierter Confounder für die der durchzuführenden Studie zugrunde liegenden Datenquelle (z. B. ein klinisches Register) vertretbar ist. Letzteres ist insbesondere davon abhängig, wie viele Confounder bereits routinemäßig in der gewählten Datenquelle erhoben werden und wie viele Confounder zusätzlich für die spezifische Fragestellung in die Datenquelle aufgenommen werden müssten.

5.3 Vorschlag für die Anforderung an eine systematische Confounderidentifikation und -erhebung

Zur Evaluation von Interventionseffekten auf Basis eines nicht randomisierten Vergleichs, sollte im Rahmen der Studienplanung a priori ein Set relevanter Confounder identifiziert werden. Hierfür wurden in GA23-02 [8] Vorschläge für die konkreten Anforderungen an die systematische Confounderidentifikation erarbeitet (bestehend aus den 3 Schritten: Identifikation, Zusammenführung und Relevanzbewertung). Das Verfahren der systematischen Confounderidentifikation ist trotz der in GA23-02 beschriebenen Möglichkeiten zur Aufwandseinsparung ressourcenintensiv. Daher sollte im vorliegenden Arbeitspapier geprüft werden, welche weiteren Möglichkeiten zur Aufwandsreduktion bestehen, sowohl für die Confounderidentifikation als auch die zugehörige Erhebung der potenziellen Confounder. Eine Reduktion der zu erhebenden und damit in der Analyse zu berücksichtigenden Confounder kann auch für Indikationen mit absehbar nur geringer zur Verfügung stehender Patientenzahl (wie z. B. seltenen Erkrankungen) erstrebenswert sein. Aus diesen Gründen wurde das Verfahren im vorliegenden Projekt mit dem Ziel der Aufwandsreduktion weiterentwickelt (siehe Abbildung 9).

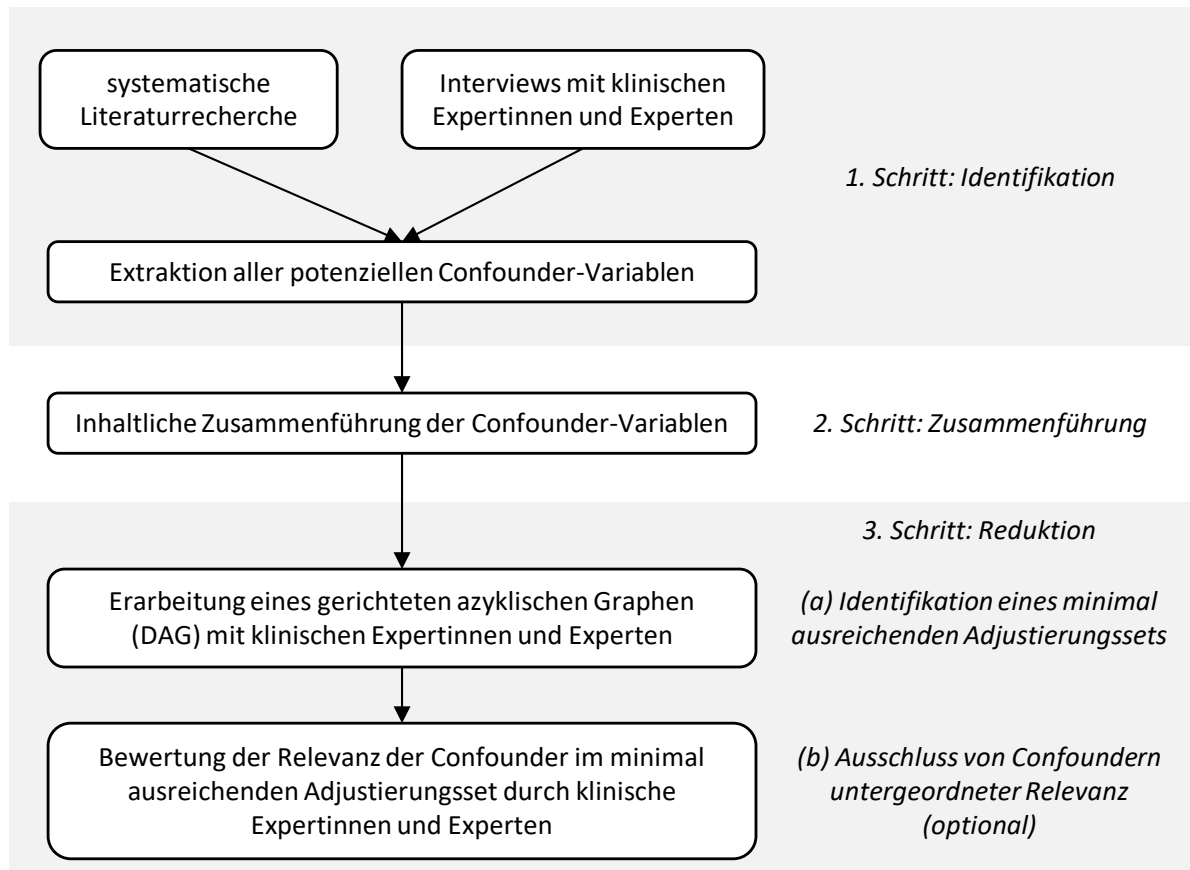


Abbildung 9: Erweiterung des in GA23-02 vorgeschlagenen Verfahrens der systematischen Confounderidentifikation um Schritt 3: Reduktion der identifizierten Confounder in 2 Teilschritten

Nach den in GA23-02 beschriebenen Schritten 1 zur Identifikation und 2 zur Zusammenführung sollte in einem 3. Schritt versucht werden, den Aufwand für die Confounder-Erhebung zu reduzieren. Hierzu sollten in einem 1. Teilschritt die angenommenen kausalen Zusammenhänge in einem DAG visualisiert werden. Dieser Schritt dient insbesondere der Identifikation des minimal ausreichenden Adjustierungssets. Gleichzeitig werden durch diesen Schritt die über die Literaturrecherche und die Interviews mit Klinikerinnen und Klinikern identifizierten und anschließend verdichteten Confounder-Sets validiert. Anschließend kann die Bewertung der Relevanz der Confounder durch klinische Expertinnen und Experten in einem 2. optionalen Teilschritt zur weiteren Reduktion des Confounder-Sets einbezogen werden.

In einem DAG werden die getroffenen Annahmen über Zusammenhänge zwischen den identifizierten Confoundern, Therapieentscheidung und Endpunkten explizit gemacht. Die Relevanz der Confounder wird nicht losgelöst, sondern im Kontext aller identifizierten, zusammengeführten Confounder betrachtet. Begründungen für weggelassene Pfeile zur Therapieentscheidung und zu Endpunkten sollten dokumentiert werden. Dies resultiert in

einem transparenten und nachvollziehbaren Vorgehen. Darüber hinaus liefert ein DAG zusammen mit dem Algorithmus zur Bestimmung minimaler Adjustierungssets eine nachvollziehbare Begründung, warum einzelne identifizierte Variablen in der Adjustierung nicht berücksichtigt – und damit auch nicht erhoben – werden müssen.

Ein weiterer Schritt zur Reduktion des Confounder-Sets sollte abhängig von der Zahl identifizierter Confounder im minimal ausreichenden Adjustierungsset bzw. der Anzahl an Patientinnen und Patienten im Anwendungsgebiet erfolgen. Die entsprechenden Cut-Offs (max. Zahl an Confoundern in Abhängigkeit der Patientenzahl) sollten a priori definiert werden. Sollte eine weitere Reduktion des Confounder-Sets notwendig sein, wird empfohlen dies anhand der von den Klinikerinnen und Klinikern angenommenen Stärke des Einflusses auf die Therapieentscheidung bzw. die Endpunkte vorzunehmen. Die Confounder sollten entsprechend der aus Therapieentscheidung und Endpunkt gemittelten Stärke ihres Einflusses nach dem Median sortiert werden und diejenigen, die unterhalb eines a priori definierten Cut-offs liegen (z. B. diejenigen, denen eine sehr geringe (1 Punkt) und geringe (2 Punkte) Bedeutung beigemessen wurde), aus dem finalen Confounder-Set entfernt werden. Hierbei sollte insbesondere transparent berichtet werden, welche Confoundern basierend auf dem gewählten Cut-off aus dem finalen Confounder-Set entfernt wurden.

Wie bereits in GA23-02 hervorgehoben, ist eine detaillierte und transparente Dokumentation der Vorgehensweise und der Ergebnisse erforderlich, um die Confounderidentifikation nachvollziehen zu können. Die Dokumentation sollte als Teil des Studienprotokolls der geplanten nicht randomisierten vergleichenden Studie veröffentlicht werden.

Die bereits in GA23-02 beschriebenen Handlungsempfehlungen, welche als Orientierungshilfe für das methodische Vorgehen bspw. im Rahmen einer AbD dienen sollen, werden nachfolgend basierend auf den Erkenntnissen des vorliegenden Projekts erweitert. Alle Empfehlungen sind auch als Checkliste in Anhang H zu finden.

5.3.1 Schritt 1: Identifikation

Zusätzlich zu den in GA23-02 beschriebenen Handlungsempfehlungen lassen sich aus den vorliegenden Untersuchungen die folgenden Empfehlungen ableiten:

Im Sinne der Aufwandsreduktion können die in hochwertigen systematischen Übersichten enthaltenen Primärpublikationen für die Extraktion potenzieller Confounder-Variablen verwendet werden, sofern durch die systematischen Übersichten die gesamte Fragestellung abgedeckt wird.

Potenzielle Confounder-Variablen sollten aus einer systematischen Literaturrecherche und aus Interviews mit klinischen Expertinnen und Experten identifiziert werden. Sofern aktuelle, qualitativ hochwertige systematische Übersichten vorliegen, die die gesamte Fragestellung

abdecken, kann die Literaturrecherche über die in den systematischen Übersichten eingeschlossenen Primärstudien erfolgen. Die systematischen Übersichten sollten hinsichtlich der Qualität der Informationsbeschaffung, z. B. mittels AMSTAR-2 (Items 1, 4 und 5) bewertet werden. Um eine sichere Abdeckung bez. Fragestellung, Studientyp und Aktualität zu erzielen, sollten als Richtwert insgesamt mindestens 2 systematische Übersichten eingeschlossen werden. Falls die identifizierten systematischen Übersichten nicht alle Aspekte der Fragestellung abdecken, ist zusätzlich eine ergänzende Suche nach Primärstudien für den zuvor nicht berücksichtigten Teilaspekt notwendig (z. B. wenn ein Studientyp für eine der beiden Interventionen fehlt).

Es sollte darauf geachtet werden, dass die systematischen Übersichten neben Kohortenstudien auch RCTs umfassen. Falls beide Studientypen nur in geringem Umfang vorliegen, z. B. bei Fragestellungen in Indikationen mit absehbar weniger umfassender Studienlage (wie bei seltenen Erkrankungen), sollten auch systematische Übersichten mit 1-armigen Studien herangezogen werden. Bei der anschließenden Extraktion sollten aus Kohortenstudien alle Faktoren, die in den statistischen Analysen zur Adjustierung genannt wurden (unabhängig von der statistischen Signifikanz), und aus RCTs und 1-armigen Studien die Patientencharakteristika berücksichtigt werden.

Der Recherchezeitpunkt nach systematischen Übersichten bzw. Primärstudien sollte nicht älter als 12 Monate sein.

Bei beiden Beispielen war das Heranziehen von Primärstudien mit einem Erscheinungsjahr bis 1 oder 2 Jahre vor der Erstellung des Referenzstandards notwendig, um nahezu alle potenziellen Confounder identifizieren zu können. Aus diesem Grund sollte die Recherche bei einem prospektiven Studiendesign zu Beginn der Datenerhebung – bzw. bei einem retrospektiven Studiendesign zum Zeitpunkt der Finalisierung des Studienprotokolls – nicht älter als 12 Monate sein. Insbesondere bei einem prospektiven Studiendesign ist eine entsprechende Aktualität unbedingt erforderlich, da die Analyse der Ergebnisse nicht zum Zeitpunkt der Confounderidentifikation, sondern ggf. erst mehrere Jahre später (nach Abschluss der Datenerhebung) durchgeführt wird.

Bei der Extraktion aus Primärpublikationen wird die Anwendung eines Sättigungskriteriums von 10 empfohlen.

Bereits in GA23-02 wurde empfohlen, ein objektives und plausibles Sättigungskriterium für die Extraktion der Primärpublikationen zu definieren. Die an dem Beispiel RRMS vorgenommenen Simulationen bestätigen das von Pufulete 2022 vorgeschlagene Sättigungskriterium. Unabhängig vom Studiendesign wird der potenzielle Informationsverlust bei einem Sättigungskriterium von 10 (d. h. 10 konsekutive Publikationen ohne neu-identifizierte Confounder-Variable) als vertretbar eingeschätzt, da in 95 % der Fälle nahezu

alle Confounder-Variablen identifiziert werden würden. Im Gegensatz dazu wird der potenzielle Informationsverlust bei einem Sättigungskriterium von 5 als zu hoch eingeschätzt.

5.3.2 Schritt 2: Zusammenführung

Die Handlungsempfehlungen für die Zusammenführung der im 1. Schritt identifizierten Variablen sind dem Arbeitspapier GA23-02 zu entnehmen.

5.3.3 Schritt 3: Reduktion

Die im Arbeitspapier GA23-02 als 3. Schritt beschriebene Relevanzbewertung wird durch die folgenden Handlungsempfehlungen aktualisiert:

Teilschritt 3a: Identifikation eines minimal ausreichenden Adjustierungssets

Für einen systematischen und transparenten Ansatz der Reduktion zu erhebender Confounder sollten unter Einbindung klinischer Expertinnen und Experten ein oder mehrere minimal ausreichende Adjustierungssets mittels gerichteter azyklischer Graphen erarbeitet werden. Es wird empfohlen, diese in einem Präsenz-Workshop mit einem Team aus mindestens 1 methodischen und 3 klinischen Expertinnen und Experten zu erarbeiten.

Durch die Erarbeitung eines DAG werden die getroffenen Annahmen für die Zusammenhänge zwischen den identifizierten Confoundern, Therapieentscheidung und Endpunkten explizit gemacht. Die Confounder werden nicht isoliert, sondern im Kontext aller identifizierten, zusammengeführten Confounder betrachtet. Im Ergebnis liefert ein DAG über die Identifikation eines minimal ausreichenden Adjustierungssets eine Begründung, warum einzelne identifizierte Variablen in der Adjustierung nicht berücksichtigt werden müssen.

Die Erarbeitung des DAG sollte gemeinsam durch ein multidisziplinäres Team bestehend aus mindestens 3 klinischen Expertinnen und Experten aus dem jeweiligen Fachgebiet und einer Methodikerin bzw. einem Methodiker mit Expertise und Erfahrung in der DAG-Erstellung erfolgen. Die Klinikerinnen und Kliniker haben idealerweise Erfahrung in der für die zu bearbeitende Fragestellung spezifischen Therapieentscheidung. Falls die zu untersuchende Intervention noch nicht lange auf dem Markt ist, sollte zumindest Erfahrung in der Anwendung der Vergleichstherapie vorhanden sein.

Es wird empfohlen den klinischen Expertinnen und Experten nicht nur die Fragestellung, für die der Graph erstellt wird, zu erläutern, sondern auch eine Einführung in die methodischen Grundzüge des Confoundings und von DAGs zu geben. Es wird empfohlen, den Workshop in Präsenz durchzuführen. Der zeitliche Umfang sollte in Abhängigkeit der Anzahl identifizierter Confounder auf 1 bis 2 Tage geplant werden.

Die Erarbeitung eines DAG kann durch eine Vorsortierung der Variablen bspw. nach inhaltlichen Aspekten (demografischen Faktoren, Life Style Faktoren, Krankheitsschwere) oder zeitlichen Aspekten (chronologisch mittels zeitlichem Raster) erleichtert werden.

Zur Erstellung des DAG wird ein pragmatischer Ansatz empfohlen: Alle Endpunkte werden gemeinsam im DAG betrachtet und es wird ein Pfeil gezeichnet, wenn ein kausaler Zusammenhang für mindestens einen Endpunkt für mindestens 1 Person angenommen wird. Im vorliegenden Projekt wurde mit einem leeren (kantenlosen) DAG gestartet. Ob mit einem leeren (kantenlosen) oder einem gesättigten DAG gestartet wird, sollte in Abhängigkeit der Anzahl identifizierter Confounder entschieden werden. Grundsätzlich birgt der Start mit einem gesättigten DAG den Vorteil, dass der Fokus auf dem Weglassen (Streichen) von Pfeilen liegt, was in einem DAG einer stärkeren Annahme entspricht. Bei einer hohen Anzahl identifizierter Confounder kann ein gesättigter DAG sehr unübersichtlich werden, dann ist es ggf. sinnvoll, mit einem leeren DAG zu starten.

Die Operationalisierung der Confounder inklusive der zeitlichen Aspekte (z. B. EDSS zu Baseline vs. EDSS-Veränderung im Vorjahr) sollte vor der Erstellung des DAG definiert werden. Gleichzeitig sollte für die zu untersuchende Fragestellung geklärt werden, ob potenzielle konkurrierende Ereignisse oder ein Selektionsknoten berücksichtigt werden sollten.

Es ist sicherzustellen, dass alle einzelnen Verbindungen zwischen den Variablen besprochen und konsentiert werden. Dieser Prozess kann durch die Dokumentation in einer Matrix erleichtert werden (siehe Anhang E.2.1). Bei fehlendem Konsens zwischen den klinischen Expertinnen und Experten hinsichtlich eines kausalen Zusammenhangs zwischen 2 Variablen sollte jeweils ein DAG mit einem Pfeil und ein DAG ohne Pfeil zwischen diesen beiden Variablen gezeichnet werden. Sofern sich unterschiedliche minimale Adjustierungssets ergeben, sollten alle Confounder, die in mindestens einem minimalen Adjustierungsset vertreten sind, erhoben werden und die verschiedenen Adjustierungssets im Rahmen von Sensitivitätsanalysen berücksichtigt werden.

Weggelassene Pfeile im Sinne eines fehlenden Einflusses auf Therapieentscheidung oder Endpunkte müssen begründet werden. Annahmen für einen fehlenden kausalen Zusammenhang zwischen den einzelnen Confoundern müssen mindestens in tabellarischer Form transparent dokumentiert werden. Im Bericht zur Confounderidentifikation muss das finale DAG bzw. die finalen DAGs inkl. Modell-Code für das DAG aus der zur Erstellung verwendeten Software und Matrix abgebildet werden.

Da die Erstellung eines DAG insbesondere auf den Annahmen der eingebundenen Klinikerinnen und Klinikern beruht, ist es wesentlich diese Annahmen für ein transparentes und nachvollziehbares Vorgehen zu dokumentieren.

Grundsätzlich ist die Annahme, dass kein Zusammenhang zwischen 2 Variablen besteht also das Weglassen eines Pfeils) eine stärkere Annahme als die Annahme, dass ein Zusammenhang besteht (also das Zeichnen eines Pfeils). Da Pfeile zur Therapieentscheidung und Endpunkt im Vergleich zu Pfeilen zwischen den einzelnen Variablen maßgeblichere Auswirkungen auf das minimale Confounder-Set haben, müssen die Begründungen für weggelassene Pfeile zur Therapieentscheidung und Endpunkt dokumentiert werden.

Plausible Gründe für das Weglassen von Pfeilen zur Therapieentscheidung sind

- 1) eine indirekte Wirkung über weitere Variablen oder
- 2) die fehlende Erhebung der Variable im Versorgungsalltag, sodass diese keinen Einfluss auf die Therapieentscheidung haben kann.

Grundsätzlich ist für diese Begründungen (für den Einfluss auf die Therapieentscheidung) die Aussage der praktizierenden Klinkerinnen und Klinikern im Konsens ausreichend. Liegt kein Konsens vor, ist (a) der Pfeil zu zeichnen oder (b) eine Sensitivitätsanalyse unter Berücksichtigung dieses Zusammenhangs durchzuführen.

Gründe für das Weglassen von Pfeilen zu den Endpunkten sind

- 1) eine (idealerweise literaturgestützte) indirekte Wirkung über weitere Variablen oder
- 2) eine aussagekräftige / literaturgestützte Begründung, dass kein Zusammenhang der betrachteten Variablen mit keinem der Endpunkte existiert.

Liegt keine ausreichende Begründung für das Weglassen eines Pfeils zum Endpunkt vor oder ist die Situation unklar, sollte grundsätzlich ein konservatives Vorgehen (im Sinne des Zeichnens eines Pfeils) gewählt werden. Zudem sollte das Fehlen von Evidenz nicht damit gleichgesetzt werden, dass kein Zusammenhang besteht. Eine einzelne Studie, die keinen statistisch signifikanten Zusammenhang zwischen den Variablen zeigt, sollte nicht zum Weglassen des Pfeils führen. Mehrere solcher Studien können aussagekräftig genug sein, um einen Pfeil zu entfernen.

Für weggelassene Pfeile zwischen den einzelnen Variablen ist im Sinne der Machbarkeit keine Dokumentation inhaltlich dezidierter Begründungen notwendig, allerdings müssen diese weggelassenen Pfeile mindestens in tabellarischer Form transparent dokumentiert werden, um einen Plausibilitätscheck zu ermöglichen.

Zur systematischen Protokollierung bereits besprochener bzw. noch offener Zusammenhänge zwischen den einzelnen Variablen wird die Verwendung einer farblich-kodierten Matrix (siehe Anhang E.2.1) empfohlen. Zur digitalen Dokumentation sollte eine angemessene Software verwendet werden, derzeit stellt DAGitty [42,43] eine geeignete Software dar.

Letztlich muss der bzw. die DAGs, der Code der verwendeten Software sowie die Matrix der dokumentierten Zusammenhänge als Teil des Studienprotokolls der geplanten nicht randomisierten vergleichenden Studie veröffentlicht werden.

Die Workshop-Teilnehmenden sollten namentlich benannt und potenzielle Interessenkonflikte offengelegt werden.

Es sollte transparent dokumentiert werden, wie viele klinische Expertinnen und Experten mit welchem Fachgebiet (bzw. Affiliation) eingebunden wurden. Die in diesem Schritt eingebundene Gruppe klinischer Expertinnen und Experten sollte im Idealfall nicht ausschließlich aus den in Schritt 1 zur Identifikation involvierten klinischen Expertinnen und Experten bestehen.

Für die Dokumentation ist es erforderlich die Workshop-Teilnehmenden (d. h. die klinischen und methodischen Expertinnen und Experten) namentlich zu benennen und potenzielle Interessenskonflikte offenzulegen (siehe Anhang H). In Anhang G ist ein Vorschlag für ein Formblatt zur Offenlegung von Beziehungen bei der Einbindung in die Confounderidentifikation dargestellt. Die Dokumentation kann z. B. in einem separaten Anhang zum Hauptteil des Berichts über die Confounderidentifikation vorgelegt werden.

Teilschritt 3b: Relevanz der Confounder – Bewertung der Stärke des Einflusses auf die Effektschätzung (optional)

Für die weitere Reduktion des finalen Confounder-Sets kann die Stärke des Einflusses sowohl auf die Therapieentscheidung als auch auf die Endpunkte einbezogen werden. Diese kann mithilfe einer Umfrage bei Klinikerinnen und Klinikern ermittelt werden.

Sollte es notwendig sein, die Confounder-Liste weiter zu reduzieren (z. B. aufgrund einer sehr geringen Patientenzahl im Anwendungsgebiet), ist zu empfehlen dies im Anschluss an die DAG-Erarbeitung anhand der angenommenen Stärke des Einflusses auf die Therapieentscheidung bzw. Endpunkte vorzunehmen. Es sollte a priori definiert werden, anhand welcher Kriterien die Entscheidung über eine weiter notwendige Reduktion getroffen wird. Diese Stärke des Einflusses der Confounder sollte mittels einer Umfrage unter klinischen Expertinnen und Experten erhoben werden.

Um einen möglichst hohen Rücklauf zu erzielen und möglichst andere Klinikerinnen und Kliniker als zuvor bei der DAG-Erstellung anzusprechen, sollte die Umfrage online erfolgen. In der Literatur gibt es keine einheitlichen Empfehlungen für die Anzahl der einzubindenden Klinikerinnen und Kliniker. Der Pool der angefragten Klinikerinnen und Kliniker sollte möglichst groß sein und sie sollten Erfahrung in der Therapieentscheidung haben, die in der Studie untersucht wird. Zudem sollte der Pool möglichst repräsentativ sein (z. B. wenn ein Arzneimittel in verschiedenen Versorgungsebenen verwendet wird, sollten diese abgebildet sein).

Es wird grundsätzlich empfohlen, den Klinikerinnen und Klinikern die zugrunde liegende Fragestellung der Umfrage ausführlich zu erläutern. Dies kann in Form eines Begleitschreibens oder in der Umfrage selbst erfolgen. Dabei muss für alle Confounder eine eindeutige Operationalisierung angegeben werden. Gleichzeitig sollte der Hinweis erfolgen, dass die Bewertung der Confounder im Kontext der gesamten Confounder-Liste erfolgen soll.

Die Abfrage des Einflusses kann mittels einer 5-Punkte-Skala von „sehr gering“ (1 Punkt) bis „sehr stark“ (5 Punkte) und sollte separat für die Therapieentscheidung und die Endpunkte erfolgen. Falls die jeweiligen Confounder die einzelnen Endpunkte unterschiedlich stark beeinflussen, sollte die Einschätzung für den Endpunkt angegeben werden, der am stärksten beeinflusst wird.

Um ein Ranking der Confounder entsprechend ihrer Einflussstärke von „sehr stark“ bis „sehr gering“ zu ermöglichen, sollten die einzelnen Bewertungen der Stärke des Einflusses auf die Therapieentscheidung bzw. auf die Endpunkte zusammengefasst werden. Dies kann beispielsweise durch die Bildung eines gewichteten Mittelwertes erreicht werden. Für die Zusammenfassung sollte eine Gewichtung zwischen Therapieentscheidung und Endpunkten definiert werden. Als pragmatisches Vorgehen wird hier eine Gewichtung von 1:1 vorgeschlagen, entsprechend des arithmetischen Mittelwertes. Anschließend können diejenigen Confounder, die unterhalb eines a priori definierten Cut-offs liegen, aus dem finalen Confounder-Set entfernt werden. Ein Cut-off bei einem Wert von ≤ 2 wird als angemessen angesehen. Dieser bedeutet, dass diejenigen Confounder, denen eine sehr geringe (1 Punkt) und geringe (2 Punkte) Bedeutung beigemessen wurde, ausgeschlossen werden. Wesentlich bei dem gesamten Vorgehen sollte eine transparente, nachvollziehbare Dokumentation des Schrittes sein inkl. der Darstellung, welche Confoundern basierend auf dem gewählten Cut-off aus dem finalen Confounder-Set entfernt wurden.

6 Diskussion

National und international spielt die Evaluation von Interventionseffekten auf Basis nicht randomisierter Vergleiche (z. B. mit versorgungsnahen Daten) eine immer größere Rolle. Einer der wichtigsten Anwendungsfälle sind vergleichende Registerstudien (wie z. B. auch im Rahmen der anwendungsbegleitenden Datenerhebung gefordert). Aber auch die Zulassung von Arzneimitteln basiert immer häufiger auf 1-armigen, nicht vergleichenden Studien. Um für letztere im Rahmen der frühen Nutzenbewertung dennoch einen Vergleich mit dem bisherigen Therapiestandard zu ermöglichen, wird dann auf externe Kontrollen, z. B. aus Registern, zurückgegriffen. In den meisten Fällen ist diese Evidenz aber nicht ausreichend, insbesondere, weil Confounder zum einen nicht systematisch identifiziert bzw. zum anderen eine ausreichende Adjustierung für Confounder nicht möglich war. Dies ist für die Aussagekraft und Interpretation von Ergebnissen nicht randomisierter Studien jedoch unabdingbar. Die Notwendigkeit einer systematischen Confounderidentifikation wird zwar vonseiten der EMA [78] ebenso wie von verschiedenen HTA-Agenturen [36,79,80] adressiert, allerdings werden bisher keine konkreten Handlungsempfehlungen zum Vorgehen beschrieben. Dies wurde bereits in mehreren Publikationen [39,81-83] adressiert und der dringende Bedarf einer praxisnahen Anleitung vorgebracht.

Angesichts dieser einheitlichen Forderungen ist es umso erstaunlicher, dass in vielen Publikationen die zur Adjustierung verwendeten Confounder nicht angegeben werden. Selbst wenn die Confounder benannt werden, erfolgt keine transparente Darstellung, wie diese identifiziert und ausgewählt wurden. So stellte eine Untersuchung zur Qualität der Berichterstattung in der aktuellen MS-Literatur fest, dass nur ca. 1 Fünftel der Studien berichtete, wie die Confounder ausgewählt wurden [38]. Auch in anderen Indikationen wurden in verschiedenen Studien verschiedene Confounder zur Adjustierung herangezogen, wobei die Methodik zur Auswahl offen blieb [84].

Die Publikation von Pufulete 2022 [12] stellt bisher den einzigen publizierten Vorschlag einer systematischen Confounderidentifikation dar, welche im Arbeitspapier GA23-02 [8] reproduziert und weiterentwickelt wurde. Die Untersuchungen zeigten jedoch auch, wie aufwendig die Confounderidentifikation für nicht randomisierte vergleichende Studien ist.

Bei Pufulete 2022 wurden 2544 Referenzen im Titel / Abstract gesichtet, 70 Studien eingeschlossen und letztendlich aus 47 Studien extrahiert. Im Arbeitspapier GA23-02 wurden 1202 Treffer in der Recherche erzielt, von denen 802 Referenzen im Titel / Abstract gesichtet und aus 80 Studien extrahiert wurde. Vor diesem Hintergrund wurde in der vorliegenden Untersuchung erstmals geprüft, ob es Möglichkeiten der Aufwandsreduktionen bei der Literaturrecherche gibt.

Erste Überlegungen zur Aufwandreduktion beinhalteten einen Rückgriff auf die Informationen in systematischen Übersichten, was sich jedoch als nicht zielführend herausstellte, da in diesen die Adjustierungsfaktoren der einzelnen Primärpublikationen selten berichtet werden. Daraufhin wurde in einem 2. Schritt geprüft, ob ein Rückgriff auf die in systematischen Übersichten eingeschlossenen Primärpublikationen ggf. ausreicht und dadurch auf eine umfassende Recherche nach Primärpublikationen verzichtet werden kann. Hier konnte in der vorliegenden Untersuchung erstmals ein Vorgehen erarbeitet werden, das den Aufwand bei der Informationsbeschaffung im Vergleich zu oben genanntem Umfang deutlich reduzieren kann. Dabei muss allerdings gewährleistet sein, dass die gesamte Fragestellung und verschiedene Studientypen durch aktuelle und hochwertige systematische Übersichten abgedeckt sind. Ist dies nicht der Fall, muss für den Teil der Fragestellung, der ggf. nicht durch die systematische Übersicht abgedeckt ist, die entsprechende Primärliteratur weiterhin direkt identifiziert werden. Die ergänzende Recherche kann je nach dem so aufwendig sein, dass nur eine geringe oder ggf. gar keine Aufwandsreduktion möglich ist.

Des Weiteren waren im Vergleich zum Vorgehen von Pufulete 2022 methodische Erweiterungen für das Beispiel MCL notwendig: Hier wurden auch 1-armige Studien mit ausgewertet. Ohne diese Erweiterung wären nur wenige potenzielle Confounder identifiziert worden, da in den systematischen Übersichten nur 1 RCT und 1 Kohortenstudie ausgewertet werden konnten. Dies bestätigt, dass die in GA23-02 empfohlene Erweiterung der Methodik von Pufulete 2022 projektspezifisch sinnvoll ist.

Neben der Reduktion des Rechercheaufwands für die Confounderidentifikation ist eine weitere Möglichkeit der Aufwandsreduktion, die Zahl der zu erhebenden Confounder zu reduzieren. Hierbei stellt die Erarbeitung eines DAG unter Einbindung von klinischen Expertinnen und Experten eine geeignete Maßnahme zur Aufwandsreduktion dar. Mithilfe eines DAG können ein oder mehrere minimal ausreichende Adjustierungssets identifiziert werden. Die in einem der Sets enthaltenen Confounder müssen in der Auswertung berücksichtigt werden, um das Risiko für Verzerrung aufgrund von Confounding zu minimieren. Für eine prospektive nicht randomisierte vergleichende Studie bedeutet dies, dass diese Confounder mit ausreichender Datenqualität erhoben werden müssen. Für eine Studie, die auf retrospektive Daten zurückgreift, müssen die Confounder mit ausreichender Qualität im Datensatz vorhanden sein. Entgegen der Vorgehensweise in einigen Studien [32,33] sollte die Verfügbarkeit eines Confounders dabei keinen Einfluss auf dessen Identifikation haben. Auch die Annahme, dass eine höhere Verfügbarkeit mit einer größeren Bedeutung einhergeht [85], ist fraglich. Sind Confounder nicht bekannt oder sind nicht messbare Confounder im Adjustierungsset enthalten, ist auch nach der Adjustierung von residuellem / unmeasured Confounding auszugehen. Diese Unsicherheiten sollten bei der Interpretation der Ergebnisse berücksichtigt werden.

Letztlich lässt sich in den zu untersuchenden Daten nicht statistisch testen, ob eine bestimmte Variable einen Confounder darstellt [86]. Ebenso lassen sich die in den verschiedenen Schritten des vorgeschlagenen Vorgehens getroffenen Annahmen nicht empirisch testen. Entsprechend können sich die Annahmen je nach eingebundener Expertengruppe unterscheiden. Als Validierung eines DAG wäre es höchstens vorstellbar, die Plausibilität der getroffenen Annahmen z. B. durch weitere Klinikerinnen und Kliniker prüfen zu lassen. Hier wäre es insbesondere hilfreich, wenn für Zusammenhänge zwischen sogenannten „core roots“, also regelhaft vorkommenden Confoundern wie z. B. Geschlecht und Alter, bereits durch Experten konsentiertere Annahmen vorliegen würden, auf welche bei einer Erarbeitung eines DAG zurückgegriffen werden könnte. Ein solcher Datensatz ist bisher nicht bekannt.

Zur Einschätzung der Unsicherheiten hinsichtlich residuellen Confoundings werden in der Literatur verschiedene alternative Verfahren, z. B. der quantitativen Bias-Analyse (umfasst den E-Value, der misst, wie stark ein nicht berücksichtigter Confounder mit Behandlung und Endpunkt assoziiert sein muss, um den Effekt zu erklären) [87-89] oder durch Verwendung von Negativ-Kontrollen [90], vorgeschlagen. Diese verfolgen das Ziel, Aussagen darüber zu erlauben, wie sicher das Ergebnis aus einer nicht randomisierten vergleichenden Studie ist. Diese Analysen können als Sensitivitätsanalysen durchgeführt werden, um den potenziellen Einfluss residuellen Confoundings abzuschätzen. Dabei ist aber zu berücksichtigen, dass diese Methoden auf zusätzlichen spezifischen Annahmen beruhen, die nicht prüfbar sind [91,92].

Vor allem bei seltenen Erkrankungen, bei denen nur wenige Patientinnen und Patienten im Anwendungsgebiet zur Verfügung stehen, ist es in der Regel nicht möglich, eine große Zahl an Confoundern in den Analysen zu berücksichtigen. In diesen Fällen kann es geboten sein – ganz ungeachtet des Ziels, den Aufwand für die Erhebung der Variablen zu reduzieren – weitere Confounder auszuschließen. Ein solcher Ausschluss von Confoundern sollte als letzter Schritt auf Basis der erwarteten Bedeutung der Confounder für die Effektschätzung erfolgen. Auch für diesen Schritt ist die Einbindung von klinischen Expertinnen und Experten unerlässlich, um den Einfluss der verschiedenen Confounder auf die Therapieentscheidung und die Endpunkte abschätzen zu können. Wird die Mehrzahl an Confoundern von den klinischen Expertinnen und Experten als wichtig eingeordnet – wie im vorliegenden Beispiel – ist jedoch keine relevante Reduktion an Confoundern möglich, ohne das Risiko eines nicht abschätzbaren Verlusts an Ergebnissicherheit einzugehen. Auch bei Indikationen mit nur geringer Patientenzahl müssen die erhöhten Anforderungen, mit denen nicht randomisierte vergleichende Analysen einhergehen, erfüllt werden.

Die in dieser Untersuchung gewonnen Erkenntnisse liefern konkrete Handlungsempfehlungen, die für die Studienplanung nicht randomisierter vergleichender Studien herangezogen werden können. Diese gelten für Studien sowohl mit prospektiver als auch mit retrospektiver Datenerhebung. Sollte sich die Qualität der Studienplanung im Sinne der

Durchführung einer systematischen Confounderidentifikation mit entsprechender Berichtsqualität aufgrund der Umsetzung der methodischen Empfehlungen zukünftig ändern, könnte dies die Confounderidentifikation deutlich erleichtern, da auf bestehende publizierte Confounderidentifikation zurückgegriffen werden kann. Voraussetzung dafür ist jedoch, dass die Confounder in einer hinreichend ähnlichen Fragestellung identifiziert wurden. In solchen Fällen kann auf die bereits durchgeführte Literaturrecherche und auf die Zusammenführung der Confounder zurückgegriffen werden. Diese sollte aber mindestens um eine Update-Recherche für die neue Fragestellung (z. B. zu anderen Wirkstoffen in der gleichen Indikation) ergänzt werden. Zusätzlich müssen die kausalen Zusammenhänge im DAG bez. der Therapieentscheidung aktualisiert und eine neue Umfrage zur Stärke des Einflusses der Confounder durchgeführt werden, da sich diese basierend auf der zugrunde liegenden Fragestellung relevant ändern können. Gleiches gilt für die im vorliegenden Projekt durchgeführte systematische Confounderidentifikation: auch wenn die zugrunde liegende Fragestellung in der aktuellen Versorgung von untergeordneter Relevanz ist, können die identifizierten Confounder als Grundlage für eine Aktualisierung entsprechend des oben skizzierten Vorgehens herangezogen werden.

Zusammenfassend zeigt sich, dass das Ziel der Aufwandsreduktion der systematischen Confounderidentifikation zumindest für die Recherche erreicht werden kann. Eine Reduktion der zu erhebenden Confounder war im vorliegenden Beispiel nur eingeschränkt möglich. Somit ist weiterhin 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, 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 [93]. Die häufige Kritik, dass RCTs keine gute externe Validität aufweisen, ließe sich durch breite Ein- und Ausschlusskriterien im Sinne einer pragmatischen RCT adressieren [94-96]. Innerhalb der RCTs stellen registerbasierte RCTs eine inhaltlich sinnvolle Möglichkeit dar, weil mit der Verwendung einer bestehenden Datenstruktur ein weiterer Effizienzgewinn einhergeht [97].

7 Fazit

Das Verfahren der systematischen Confounderidentifikation ist ressourcenintensiv. Jedoch bestehen Möglichkeiten zur Aufwandsreduktion sowohl für die Confounderidentifikation als auch ggf. die zugehörige Erhebung der Confounder.

Anhand des RRMS-Beispiels konnte gezeigt werden, dass die Informationsbeschaffung für die Confounderidentifikation über systematische Übersichten im Vergleich zu einer Recherche nach Primärstudien grundsätzlich zeitsparender und eine Möglichkeit zur deutlichen Aufwandsreduktion sein kann. Hier ist jedoch zu beachten, dass dies nur in Situationen möglich ist, in denen über systematische Übersichten die gesamte Fragestellung abgedeckt werden kann. Die entsprechende Recherche für die Confounderidentifikation sollte bei einem prospektiven Studiendesign zu Beginn der Datenerhebung – bzw. bei einem retrospektiven Studiendesign zum Zeitpunkt der Finalisierung des Studienprotokolls – nicht älter als 12 Monate sein.

Als geeignete Möglichkeiten zur Aufwandsreduktion bei der Erhebung von Confoundern erwiesen sich:

- die Erarbeitung eines DAG mit klinischen Expertinnen und Experten und
- die anschließende Bewertung der Relevanz der Confounder durch klinische Expertinnen und Experten.

Mithilfe von DAGs kann systematisch und transparent ermittelt werden, für welche Teilmenge der identifizierten Confounder mindestens adjustiert werden muss, um alle identifizierten Confounder adäquat in der Auswertung zu berücksichtigen. Auch wenn die für die DAG-Erarbeitung getroffenen Annahmen der Expertengruppe nicht überprüfbar sind, liefern DAGs dennoch eine transparente und nachvollziehbare Begründung für das Weglassen einzelner Confounder.

Sollte eine weitere Reduktion der Confounder-Liste notwendig sein (z. B. aufgrund einer hohen Zahl identifizierter Confounder bei nur geringer für die Analysen zur Verfügung stehender Fallzahl), wird empfohlen, diese Reduktion anhand der von klinischen Expertinnen und Experten angenommenen Stärke des Einflusses auf die Therapieentscheidung bzw. die Endpunkte vorzunehmen. Alle Confounder, die im minimal ausreichenden Adjustierungsset enthalten sind, und denen nur ein untergeordneter Einfluss beigemessen wird, können aus dem finalen Confounder-Set entfernt werden. Wird die Mehrzahl an Confoundern als wichtig eingeordnet, ist keine relevante Aufwandsreduktion möglich ohne das Risiko eines nicht abschätzbaren Verlusts an Ergebnissicherheit einzugehen.

Die beschriebenen Maßnahmen zur Aufwandsreduktion können in das im Arbeitspapier GA23-02 vorgeschlagene Verfahren zur systematischen Confounderidentifikation integriert werden. Dieses setzt sich nun aus den folgenden Schritten zusammen:

- 3) Confounderidentifikation bestehend aus einer systematischen Literaturrecherche und aus Interviews mit klinischen Expertinnen und Experten,
- 4) Zusammenführung von Confounder-Variablen, welche ein inhaltlich gleiches Konstrukt messen oder sich inhaltlich überlappen,
- 5) Reduktion des zuvor identifizierten Confounder-Sets in 2 Teilschritten:
 - a) Identifikation eines minimal ausreichenden Adjustierungssets mithilfe eines DAG und unter Einbindung klinischer Expertinnen und Experten
 - b) Ausschluss von Confoundern untergeordneter Relevanz auf Basis der Einschätzung durch klinische Expertinnen und Experten (optional)

Die Durchführung dieser Schritte ist dabei vollumfänglich zu dokumentieren.

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 eine nahezu unverzerrte Effektschätzung durch die im minimal ausreichenden Adjustierungsset identifizierten und durch die Bewertung der Relevanz ausgewählten Confounder erreicht werden kann. 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 für die keine Confounderidentifikation und -kontrolle erforderlich ist. Registerbasierte RCTs verbinden dabei verschiedene Vorteile, erstens diejenigen der Randomisierung und zweitens diejenigen der Nutzung versorgungsnaher Daten in einer bestehenden Infrastruktur.

8 Literatur

1. 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>.
2. 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/>.
3. 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>.
4. Rhodes AE, Lin E, Streiner DL. Confronting the confounders: the meaning, detection, and treatment of confounders in research. *Can J Psychiatry* 1999; 44(2): 175-179. <https://doi.org/10.1177/070674379904400209>.
5. 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>.
6. Hernán MA, Robins JM. Causal Inference: What If [online]. 2024 [Zugriff: 14.09.2026]. URL: https://miguelhernan.org/s/hernanrobins_WhatIf_19aug26.pdf.
7. Bashir NZ, Kebschull M, Burgess S et al. Causal Inference in Clinical Research: A Primer. *J Periodontal Res* 2026. <https://doi.org/10.1111/jre.70142>.
8. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Systematische Confounderidentifikation in der Indikation schubförmig remittierende multiple Sklerose (RRMS); Arbeitspapier [online]. 2025 [Zugriff: 18.09.2025]. URL: <https://doi.org/10.60584/GA23-02>.
9. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Wissenschaftliche Ausarbeitung ausgewählter Aspekte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach § 35a SGB V; vorläufiger Rapid Report [online]. 2025 [Zugriff: 18.09.2025]. URL: https://www.iqwig.de/download/a25-13_generierung-und-auswertung-versorgungsnaher-daten-zum-zwecke-der-nutzenbewertung_vorlaeufiger-rapid-report_v1-0.pdf.
10. Drescher K, Timm J, Jockel KH. The design of case-control studies: the effect of confounding on sample size requirements. *Stat Med* 1990; 9(7): 765-776. <https://doi.org/10.1002/sim.4780090706>.
11. Haneuse S, Schildcrout J, Gillen D. A two-stage strategy to accommodate general patterns of confounding in the design of observational studies. *Biostatistics* 2012; 13(2): 274-288. <https://doi.org/10.1093/biostatistics/kxr044>.

12. 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>.
13. Gemeinsamer Bundesausschuss. Anwendungsbegleitende Datenerhebung und Beschränkung der Versorgungsbefugnis: Brexucabtagen-Autoleucel/Autologe Anti-CD19-transduzierte CD3+ Zellen – Mantelzelllymphom [online]. 2022 [Zugriff: 06.10.2025]. URL: <https://www.g-ba.de/anwendungsbegleitende-datenerhebung-verfahren/5/>.
14. Gemeinsamer Bundesausschuss. Beschluss des Gemeinsamen Bundesausschusses über eine Feststellung im Verfahren der anwendungsbegleitenden Datenerhebung und von Auswertungen nach § 35a Absatz 3b des Fünften Buches Sozialgesetzbuch (SGB V): Autologe Anti-CD19-transduzierte CD3+ Zellen (rezidiertes oder refraktäres Mantelzell-Lymphom) – Vorlage von Studienprotokoll und Statistischem Analyseplan [online]. 2023 [Zugriff: 02.05.2025]. URL: https://www.g-ba.de/downloads/39-261-5904/2023-03-16_AM-RL-XII_BrexCel_2021-AbD-008_Feststellung.pdf.
15. Nussbaumer-Streit B, Ellen M, Klerings I et al. Resource use during systematic review production varies widely: a scoping review. *J Clin Epidemiol* 2021; 139: 287-296. <https://doi.org/10.1016/j.jclinepi.2021.05.019>.
16. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Alemtuzumab, Cladribin, Dimethylfumarat, Fingolimod, Natalizumab, Ocrelizumab, Ofatumumab, Ozanimod, Ponesimod und Teriflunomid zur Behandlung Erwachsener mit hochaktiver schubförmig remittierender multipler Sklerose; Abschlussbericht [online]. 2023 [Zugriff: 29.08.2026]. URL: https://www.iqwig.de/download/a20-60_multiple-sklerose_abschlussbericht_v1-0.pdf.
17. Landmeyer NC, Burkner PC, Wiendl H et al. Disease-modifying treatments and cognition in relapsing-remitting multiple sclerosis: A meta-analysis. *Neurology* 2020; 94(22): e2373-e2383. <https://doi.org/10.1212/wnl.00000000000009522>.
18. Li H, Hu F, Zhang Y et al. Comparative efficacy and acceptability of disease-modifying therapies in patients with relapsing-remitting multiple sclerosis: a systematic review and network meta-analysis. *J Neurol* 2020; 267(12): 3489-3498. <https://doi.org/10.1007/s00415-019-09395-w>.
19. Liu Z, Liao Q, Wen H et al. Disease modifying therapies in relapsing-remitting multiple sclerosis: A systematic review and network meta-analysis. *Autoimmun Rev* 2021; 20(6): 102826. <https://doi.org/10.1016/j.autrev.2021.102826>.
20. Gonzalez-Lorenzo M, Ridley B, Minozzi S et al. Immunomodulators and immunosuppressants for relapsing-remitting multiple sclerosis: a network meta-analysis. *Cochrane Database Syst Rev* 2024; 1: CD011381. <https://doi.org/10.1002/14651858.CD011381.pub3>.

21. Tramacere I, Virgili G, Perduca V et al. Adverse effects of immunotherapies for multiple sclerosis: a network meta-analysis. *Cochrane Database Syst Rev* 2023; 11: CD012186. <https://doi.org/10.1002/14651858.CD012186.pub2>.
22. Prosperini L, Haggiag S, Ruggieri S et al. Dimethyl Fumarate or Teriflunomide for Relapsing-Remitting Multiple Sclerosis: A Meta-analysis of Post-marketing Studies. *Neurotherapeutics* 2023; 20(5): 1275-1283. <https://doi.org/10.1007/s13311-023-01416-x>.
23. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Evaluation zur Relevanz von Embase für die Suche nach Arzneimittelstudien; Arbeitspapier [online]. 2025 [Zugriff: 10.12.2025]. URL: <https://doi.org/10.60584/GA24-04>.
24. Roufarshbaf M, Javeri M, Akbari V et al. Efficacy and safety of ibrutinib in mantle cell lymphoma: A systematic review and meta-analysis. *Daru: Journal of Faculty of Pharmacy, Tehran University of Medical Sciences* 2022; 30(2): 367-378. <https://doi.org/10.1007/s40199-022-00444-w>.
25. Monga N, Garside J, Davids MS et al. Systematic Literature Review of Economic Evaluations, Costs/Resource Use, and Quality of Life in Patients with Mantle Cell Lymphoma. *Pharmacoeconomics Open* 2021; 5(2): 175-186. <https://doi.org/10.1007/s41669-020-00231-w>.
26. Sancho JM, Sorigue M, Rubio-Azpeitia E. Real-World Evidence of Relapsed/Refractory Mantle Cell Lymphoma Patients and Treatments: A Systematic Review. *J Blood Med* 2024; 15: 239-254. <https://doi.org/10.2147/jbm.S463946>.
27. National Institute for Health and Care Excellence. Brexucabtagene autoleucl for treating relapsed or refractory mantle cell lymphoma; Technology appraisal guidance TA677 [online]. 2021 [Zugriff: 02.09.2026]. URL: <https://www.nice.org.uk/guidance/ta677>.
28. J JW, Wade SW, Itani T et al. Unmet needs in relapsed/refractory mantle cell lymphoma (r/r MCL) post-covalent Bruton tyrosine kinase inhibitor (BTKi): a systematic literature review and meta-analysis. *Leuk Lymphoma* 2024: 1-14. <https://doi.org/10.1080/10428194.2024.2369653>.
29. Visco C, Di Rocco A, Evangelista A et al. Outcomes in first relapsed-refractory younger patients with mantle cell lymphoma: results from the MANTLE-FIRST study. *Leukemia* 2020; 35(3): 787-795. <https://doi.org/10.1038/s41375-020-01013-3>.
30. Shea BJ, Reeves BC, Wells G et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017; 358: j4008. <https://doi.org/10.1136/bmj.j4008>.

31. 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 (DOI: 10.60584/GA23-02) [online]. [Zugriff: 29.08.2026]. URL: https://www.iqwig.de/download/ga23-02_supplement-zur-externen-expertise_anhang-f.xlsx.
32. Abrisqueta P, Jimenez-Ubieto A, Serna A et al. Prognostic factors and effect modifiers in patients with relapsed or refractory follicular lymphoma who failed at least two lines of therapy: a systematic literature and expert clinical review. *Ann Hematol* 2025; 104(9): 4357-4367. <https://doi.org/10.1007/s00277-025-06575-9>.
33. von Tresckow B, Abrisqueta P, Zamanillo I et al. Prognostic Factors and Effect Modifiers in Patients With Relapse or Refractory Diffuse Large B-Cell Lymphoma After Two Lines of Therapy: A Systematic Literature and Expert Clinical Review. *Eur J Haematol* 2025; 115(2): 104-116. <https://doi.org/10.1111/ejh.14423>.
34. Kumar S, Leleu X, Weisel KC et al. Ranking the Importance of Prognostic Factors for Relapsed/Refractory Multiple Myeloma: International Physician Panel Consensus Following a Systematic Literature Review. *Clin Lymphoma Myeloma Leuk* 2025; 25(10): 730-738.e11. <https://doi.org/10.1016/j.clml.2025.03.013>.
35. European Medicines Agency. ICH M14 Guideline on general principles on planning, designing, analysing, and reporting of non-interventional studies that utilise Real-World Data for safety assessment of medicines, Step 5, EMA/CHMP/ICH/155061/2024 [online]. 2025 [Zugriff: 23.12.2025]. URL: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-m14-guideline-general-principles-planning-designing-analysing-reporting-non-interventional-studies-utilise-real-world-data-safety-assessment-medicines-step-5_en.pdf.
36. National Institute for Health and Care Excellence. NICE real-world evidence framework (ECD9) [online]. 2025 [Zugriff: 19.12.2025]. URL: <https://www.nice.org.uk/guidance/ecd9/resources/nice-realworld-evidence-framework-pdf-1124020816837>.
37. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Wissenschaftliche Ausarbeitung ausgewählter Aspekte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach § 35a SGB V; Dokumentation der Anhörung zum vorläufigen Rapid Report [online]. 2025 [Zugriff: 01.12.2025]. URL: https://www.iqwig.de/download/a25-13_generierung-und-auswertung-versorgungsnaher-daten-zum-zwecke-der-nutzenbewertung_da-vorlaeufiger-rapid-report_v1-0.pdf.
38. Karim ME, Pellegrini F, Platt RW et al. The use and quality of reporting of propensity score methods in multiple sclerosis literature: A review. *Mult Scler* 2022; 28(9): 1317-1323. <https://doi.org/10.1177/1352458520972557>.

39. Neumann MA, Jost J, Riou S et al. Covariate selection and adjustment for efficacy and safety endpoints in indirect comparative effectiveness analyses of CAR-T-cell therapies for large B-cell lymphoma: a systematic review. *J Comp Eff Res* 2026; e260033. <https://doi.org/10.57264/cer-2026-0033>.
40. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden; Version 8.0 [online]. 2025 [Zugriff: 06.01.2026]. URL: https://doi.org/10.60584/Allgemeine-Methoden_V8.0.
41. Sterne J, Mathur MB, Elbers R et al. The Risk Of Bias In Non-randomized Studies – of Interventions, Version 2 (ROBINS-I V2) assessment tool (for follow-up studies) [online]. 2025 [Zugriff: 08.06.2026]. URL: <https://www.riskofbias.info/welcome/robins-i-v2>.
42. Textor J, Hardt J, Knuppel S. DAGitty: a graphical tool for analyzing causal diagrams. *Epidemiology* 2011; 22(5): 745. <https://doi.org/10.1097/EDE.0b013e318225c2be>.
43. Textor J, van der Zander B, Gilthorpe MS et al. Robust causal inference using directed acyclic graphs: the R package ‘dagitty’. *Int J Epidemiol* 2016; 45(6): 1887-1894. <https://doi.org/10.1093/ije/dyw341>.
44. Braitmaier M, Kollhorst B, Heinig M et al. Effectiveness of Mammography Screening on Breast Cancer Mortality - A Study Protocol for Emulation of Target Trials Using German Health Claims Data. *Clin Epidemiol* 2022; 14: 1293-1303. <https://doi.org/10.2147/CLEP.S376107>.
45. Mountford S, Kahn M, Balakrishnan P et al. Correlation and differences of patient-reported outcomes vs. Likert-Rating of MS symptoms in a real-world cohort using a digital patient app. *Digit Health* 2023; 9: 20552076231173520. <https://doi.org/10.1177/20552076231173520>.
46. Wellhagen GJ, Ueckert S, Nielsen E et al. Predictors of Ability to Work in Multiple Sclerosis. *CPT Pharmacometrics Syst Pharmacol* 2025; 14(12): 2244-2251. <https://doi.org/10.1002/psp4.70143>.
47. Witte J, Didelez V. Covariate selection strategies for causal inference: Classification and comparison. *Biom J* 2019; 61(5): 1270-1289. <https://doi.org/10.1002/bimj.201700294>.
48. Bhattacharya J, Vogt W. Do Instrumental Variables Belong in Propensity Scores? [online]. 2009 [Zugriff: 29.06.2026]. URL: <https://www.nber.org/papers/t0343>.
49. Wooldridge JM. Should instrumental variables be used as matching variables? *Research in Economics* 2016; 70(2): 232-237. <https://doi.org/10.1016/j.rie.2016.01.001>.
50. Rubin DB, Thomas N. Matching using estimated propensity scores: relating theory to practice. *Biometrics* 1996; 52(1): 249-264.

51. 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>.
52. Bergstra SA, Sepriano A, Ramiro S et al. Three handy tips and a practical guide to improve your propensity score models. *RMD Open* 2019; 5(1): e000953. <https://doi.org/10.1136/rmdopen-2019-000953>.
53. Zawadzki RS, Grill JD, Gillen DL et al. 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>.
54. VanderWeele TJ. Principles of confounder selection. *Eur J Epidemiol* 2019; 34(3): 211-219. <https://doi.org/10.1007/s10654-019-00494-6>.
55. Inoue K, Sakamaki K, Komukai S et al. Methodological Tutorial Series for Epidemiological Studies: Confounder Selection and Sensitivity Analyses to Unmeasured Confounding From Epidemiological and Statistical Perspectives. *J Epidemiol* 2025; 35(1): 3-10. <https://doi.org/10.2188/jea.JE20240082>.
56. Myers JA, Rassen JA, Gagne JJ et al. Effects of adjusting for instrumental variables on bias and precision of effect estimates. *Am J Epidemiol* 2011; 174(11): 1213-1222. <https://doi.org/10.1093/aje/kwr364>.
57. Rajkumar A, Ms V. In reference to 'Directed acyclic graphs for clinical research: a tutorial'. *J Minim Invasive Surg* 2023; 26(4): 222-223. <https://doi.org/10.7602/jmis.2023.26.4.222>.
58. Pearce N, Lawlor DA. Causal inference-so much more than statistics. *Int J Epidemiol* 2016; 45(6): 1895-1903. <https://doi.org/10.1093/ije/dyw328>.
59. Tennant PWG, Murray EJ, Arnold KF et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. *Int J Epidemiol* 2021; 50(2): 620-632. <https://doi.org/10.1093/ije/dyaa213>.
60. Van der Zander B, Textor J, Liśkiewicz M. Efficiently Finding Conditional Instruments for Causal Inference [online]. 2015 [Zugriff: 11.09.2026]. URL: <https://www.ijcai.org/Proceedings/15/Papers/457.pdf>.
61. Van der Zander B, Liśkiewicz M. Separators and Adjustment Sets in Markov Equivalent DAGs [online]. 2016. URL: <https://cdn.aaai.org/ojs/10424/10424-13-13952-1-2-20201228.pdf>.
62. Van der Zander B, Liśkiewicz M, Textor J. Constructing Separators and Adjustment Sets in Ancestral Graphs. [online]. 2014 [Zugriff: 11.09.2026]. URL: <https://auai.org/uai2014/proceedings/individuals/209.pdf>.

63. Johannes Textor, Maciej Liśkiewicz. Adjustment Criteria in Causal Diagrams: An Algorithmic Perspective [online]. 2011 [Zugriff: 11.09.2026]. URL: https://www.tcs.uni-luebeck.de/downloads/papers/2011/Textor_Liskiewicz_Adjustment_Criteria_in_Causal_Diagrams_An_Algorithmic_Perspective.pdf.
64. Poppe L, Steen J, Loh WW et al. How to develop causal directed acyclic graphs for observational health research: a scoping review. Health Psychol Rev 2025; 19(1): 45-65. <https://doi.org/10.1080/17437199.2024.2402809>.
65. Digitale JC, Martin JN, Glymour MM. Tutorial on directed acyclic graphs. J Clin Epidemiol 2022; 142: 264-267. <https://doi.org/10.1016/j.jclinepi.2021.08.001>.
66. Tonnies T, Kahl S, Kuss O. Collider Bias in Observational Studies. Dtsch Arztebl Int 2022; 119(7): 107-122. <https://doi.org/10.3238/arztebl.m2022.0076>.
67. Ferguson KD, McCann M, Katikireddi SV et al. Evidence synthesis for constructing directed acyclic graphs (ESC-DAGs): a novel and systematic method for building directed acyclic graphs. Int J Epidemiol 2020; 49(1): 322-329. <https://doi.org/10.1093/ije/dyz150>.
68. Rodrigues D, Kreif N, Lawrence-Jones A et al. Reflection on modern methods: constructing directed acyclic graphs (DAGs) with domain experts for health services research. Int J Epidemiol 2022; 51(4): 1339-1348. <https://doi.org/10.1093/ije/dyac135>.
69. Petersen AH, Ekstrom CT, Spirtes P et al. Constructing Causal Life-Course Models: Comparative Study of Data-Driven and Theory-Driven Approaches. Am J Epidemiol 2023; 192(11): 1917-1927. <https://doi.org/10.1093/aje/kwad144>.
70. Swarbrick CJ, Blake HA, Martin P et al. Directed acyclic graphs to minimise bias and optimise causal inference in SNAP-3: an observational cohort study of frailty, multimorbidity, and delirium in older surgical patients. Br J Anaesth 2025; 135(1): 177-187. <https://doi.org/10.1016/j.bja.2025.04.027>.
71. Ben Noon G, Vigiser I, Shiner T et al. Reinforcing the evidence of oligoclonal bands as a prognostic factor in patients with Multiple sclerosis. Mult Scler Relat Disord 2021; 56: 103220. <https://doi.org/10.1016/j.msard.2021.103220>.
72. Maciejowska-Szydło K, Puz P. Early Prognostic Factors in Multiple Sclerosis: Clinical and Therapeutic Implications. Medicina (Kaunas) 2026; 62(3). <https://doi.org/10.3390/medicina62030475>.
73. Samadzade U, Alizada S, Zengin ES et al. Revisiting the diagnostic marker: Oligoclonal band positivity is not associated with early disability or cognitive performance in relapsing-remitting multiple sclerosis. Mult Scler Relat Disord 2025; 104: 106727. <https://doi.org/10.1016/j.msard.2025.106727>.
74. Siritho S, Freedman MS. The prognostic significance of cerebrospinal fluid in multiple sclerosis. J Neurol Sci 2009; 279(1-2): 21-25. <https://doi.org/10.1016/j.jns.2008.12.029>.

75. Hemmer B, Gehring K. Diagnose und Therapie der Multiplen Sklerose, Neuromyelitis-optica-Spektrum-Erkrankung (NMOSD) und MOG-IgG-assoziierten Erkrankung (MOGAD); S2k-Leitlinie [online]. 2026 [Zugriff: 19.05.2026]. URL: https://www.dgn.org/cms-content/030050_Living_Guideline_MS_V9_2026_1777557729023.pdf.
76. Cinelli C, Forney A, Pearl J. A Crash Course in Good and Bad Controls. *Sociological Methods & Research* 2022; 53(3): 1071-1104. <https://doi.org/10.1177/00491241221099552>.
77. Stokes T, Steele R, Shrier I. Causal simulation experiments: Lessons from bias amplification. *Stat Methods Med Res* 2022; 31(1): 3-46. <https://doi.org/10.1177/0962280221995963>.
78. European Medicines Agency. Reflection paper on use of real-world data in non-interventional studies to generate real-world evidence for regulatory purposes, EMA/99865/2025 [online]. 2025 [Zugriff: 23.12.2025]. URL: https://www.ema.europa.eu/en/documents/other/reflection-paper-use-real-world-data-non-interventional-studies-generate-real-world-evidence-regulatory-purposes_en.pdf.
79. 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: 02.09.2025]. URL: https://www.iqwig.de/download/a19-43_versorgungснаhe-daten-zum-zwecke-der-nutzenbewertung_rapid-report_v1-1.pdf.
80. 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>.
81. Freitag A, Gurskyte L, Sarri G. Increasing transparency in indirect treatment comparisons: is selecting effect modifiers the missing part of the puzzle? A review of methodological approaches and critical considerations. *J Comp Eff Res* 2023; 12(10): e230046. <https://doi.org/10.57264/cer-2023-0046>.
82. Capkun G, Corry S, Dowling O et al. Can we use existing guidance to support the development of robust real-world evidence for health technology assessment/payer decision-making? *Int J Technol Assess Health Care* 2022; 38(1): e79. <https://doi.org/10.1017/S0266462322000605>.
83. Jaksa A, Wu J, Jonsson P et al. Organized structure of real-world evidence best practices: moving from fragmented recommendations to comprehensive guidance. *J Comp Eff Res* 2021; 10(9): 711-731. <https://doi.org/10.2217/cer-2020-0228>.

84. Riou S, Rungaldier S, Mahlich J. Identification of Adjustment Variables in Indirect Comparisons: A Rapid Review of CAR-T Therapies for Diffuse Large B-Cell Lymphoma. *Cancers (Basel)* 2025; 17(8). <https://doi.org/10.3390/cancers17081335>.
85. Chen J, Zhou X, Chen R et al. Development of a Core Confounder Set for Real-World Essential Hypertension Studies. *Cardiology* 2025: 1-12. <https://doi.org/10.1159/000544860>.
86. 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.0b013e3181dbebe3>.
87. Brown PJ, Hunnicutt JN, Sanni Ali M et al. Quantifying possible bias in clinical and epidemiological studies with quantitative bias analysis: common approaches and limitations. *BMJ* 2024; 385: e076365. <https://doi.org/10.1136/bmj-2023-076365>.
88. Ioannidis JPA, Tan YJ, Blum MR. Limitations and Misinterpretations of E-Values for Sensitivity Analyses of Observational Studies. *Ann Intern Med* 2019; 170(2): 108-111. <https://doi.org/10.7326/M18-2159>.
89. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann Intern Med* 2017; 167(4): 268-274. <https://doi.org/10.7326/M16-2607>.
90. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology* 2010; 21(3): 383-388. <https://doi.org/10.1097/EDE.0b013e3181d61eeb>.
91. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Wissenschaftliche Ausarbeitung ausgewählter Aspekte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach § 35a SGB V; Rapid Report [online]. 2025 [Zugriff: 01.12.2025]. URL: <https://doi.org/10.60584/A25-13>.
92. Hoffmann S, Morris T, Herrmann M et al. Using routinely collected data for research purposes: challenges and mitigation strategies. *BMJ* 2026; 393: e087812. <https://doi.org/10.1136/bmj-2025-087812>.
93. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Evidenz zu Orphan Drugs; Arbeitspapier [online]. 2021 [Zugriff: 30.08.2026]. URL: https://www.iqwig.de/download/ga21-01_evidenz-zu-orphan-drugs_arbeitspapier_v1-0.pdf.
94. Mathes T, Buehn S, Prengel P et al. Registry-based randomized controlled trials merged the strength of randomized controlled trials 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>.
95. Lange S, Lauterberg J. Pragmatische randomisierte Studien mit Fokus auf Registerbasierung. *Präv Gesundheitsf* 2022. <https://doi.org/10.1007/s11553-022-00974-w>.

96. Ford I, Norrie J. Pragmatic Trials. N Engl J Med 2016; 375(5): 454-463.
<https://doi.org/10.1056/NEJMra1510059>.
97. 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>.
98. Wong SS, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455.
99. Waffenschmidt S, Navarro-Ruan T, Hobson N et al. Development and validation of study filters for identifying controlled non-randomized studies in PubMed and Ovid MEDLINE. Res Synth Methods 2020; 11(5): 617-626. <https://doi.org/10.1002/jrsm.1425>.
100. Whiting P, Savovic J, Higgins JP et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol 2016; 69: 225-234.
<https://doi.org/10.1016/j.jclinepi.2015.06.005>.

Anhang A Suchstrategien

A.1 Recherchen zur Confounderidentifikation über systematische Übersichten und die darin enthaltenen Primärstudien

A.1.1 RRMS

A.1.1.1 Suche nach systematischen Übersichten

1. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) ALL 1946 to June 28, 2024

Es wurde folgender Filter übernommen:

- Systematische Übersicht: Wong [98] – High specificity strategy (adaptiert)

#	Searches
1	exp "multiple sclerosis, relapsing-remitting"/
2	((multiple sclerosis or MS) and (relaps* or remitting*)).ti,ab.
3	or/1-2
4	exp "Dimethyl fumarate"/ or exp "Glatiramer acetate"/
5	(Dimethyl fumarat* or Dimethylfumarat* or Tecfidera or Copaxone or Glatiramer*).ti,ab.
6	or/4-5
7	Cochrane database of systematic reviews.jn.
8	(search or MEDLINE or systematic review).tw.
9	meta analysis.pt.
10	or/7-9
11	10 not (exp animals/ not humans.sh.)
12	and/3,6,11
13	12 and (english or german or multilingual or undetermined).lg.
14	limit 13 to yr="2019 -Current"

2. International HTA Database

Suchoberfläche: INAHTA

#	Searches
1	"Multiple Sclerosis Relapsing-Remitting"[mhe]
2	(multiple sclerosis or MS) and (relaps* or remitting*)
3	((multiple sclerosis or MS) and (relaps* or remitting*)) OR ("Multiple Sclerosis Relapsing-Remitting"[mhe])
4	"Dimethyl Fumarate"[mhe]
5	"Glatiramer Acetate"[mhe]
6	Dimethyl fumarat* or Dimethylfumarat* or Tecfidera or Copaxone or Glatiramer*
7	(Dimethyl fumarat* or Dimethylfumarat* or Tecfidera or Copaxone or Glatiramer*) OR ("Glatiramer Acetate"[mhe]) OR ("Dimethyl Fumarate"[mhe])
8	((Dimethyl fumarat* or Dimethylfumarat* or Tecfidera or Copaxone or Glatiramer*) OR ("Glatiramer Acetate"[mhe]) OR ("Dimethyl Fumarate"[mhe])) AND (((multiple sclerosis or MS) and (relaps* or remitting*)) OR ("Multiple Sclerosis Relapsing-Remitting"[mhe]))
9	* FROM 2019 TO 2025
10	(* FROM 2019 TO 2025) AND (((Dimethyl fumarat* or Dimethylfumarat* or Tecfidera or Copaxone or Glatiramer*) OR ("Glatiramer Acetate"[mhe]) OR ("Dimethyl Fumarate"[mhe])) AND (((multiple sclerosis or MS) and (relaps* or remitting*)) OR ("Multiple Sclerosis Relapsing-Remitting"[mhe])))

A.1.1.2 Suche nach Primärstudien (zur Aufwandsschätzung)

MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) ALL <1946 to March 30, 2026>

Es wurde folgender Filter übernommen:

- Non-RCT: Search filter with best specificity for controlled NRS (Ovid MEDLINE, adapted from PubMed) [99]

#	Searches
1	exp "multiple sclerosis, relapsing-remitting"/
2	((multiple sclerosis or MS) and (relaps* or remitting*)).ti,ab.
3	or/1-2
4	exp "Glatiramer acetate"/
5	(Copaxone or Glatiramer*).ti,ab.
6	or/4-5
7	and/3,6
8	exp epidemiologic studies/ or exp clinical trial/ or comparative study/
9	((control and study) or program).mp.
10	or/8-9
11	(animals/ not humans/) or comment/ or editorial/ or exp review/ or meta analysis/ or consensus/ or exp guideline/
12	hi.fs. or case report.mp.
13	or/11-12
14	10 not 13
15	7 and 14
16	..l/ 15 yr=1946-2023

A.1.2 MCL

A.1.2.1 Suche nach systematischen Übersichten

1. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) ALL 1946 to August 20, 2024

Es wurde folgender Filter übernommen:

- Systematische Übersicht: Wong [98] – High specificity strategy (adaptiert)

#	Searches
1	(mantle and cell and lymphoma).mp.
2	Cochrane database of systematic reviews.jn.
3	(search or MEDLINE or systematic review).tw.
4	meta analysis.pt.
5	or/2-4
6	5 not (exp animals/ not humans.sh.)
7	and/1,6
8	7 and (english or german or multilingual or undetermined).lg.
9	limit 8 to yr="2019 -Current"

2. International HTA Database

Suchoberfläche: INAHTA

#	Searches
1	mantle AND cell AND lymphoma

Anhang B Bewertung der Qualität der berücksichtigten systematischen Übersichten

Die Bewertung der Qualität der Informationsbeschaffung der systematischen Übersichten erfolgte in Anlehnung an AMSTAR-2, Items 1, 4 und 5 [30].

B.1 Beispiel RRMS

Tabelle 19: Bewertung der Informationsbeschaffung zur systematischen Übersicht Prosperini 2023 [22]

1. Did the research questions and inclusion criteria for the review include the components of PICO?	✓ yes no
4. Did the review authors use a comprehensive literature search strategy?	yes ✓ partial yes no
5. Did the review authors perform study selection in duplicate?	✓ yes ^a no
a. Nur 1 Reviewer selektierte initial anhand der Titel; danach erfolgte der weitere Screeningprozess durch 2 Personen unabhängig voneinander. Daher wurde der gesamte Prozess mit „ja“ bewertet.	

Tabelle 20: Bewertung der Informationsbeschaffung zur systematischen Übersicht Gonzalez-Lorenzo 2024 [20]

1. Did the research questions and inclusion criteria for the review include the components of PICO?	✓ yes no
4. Did the review authors use a comprehensive literature search strategy?	✓ yes partial yes no
5. Did the review authors perform study selection in duplicate?	✓ yes no

Tabelle 21: Bewertung der Informationsbeschaffung zur systematischen Übersicht Tramacere 2023 [21]

1. Did the research questions and inclusion criteria for the review include the components of PICO?	✓ yes no
4. Did the review authors use a comprehensive literature search strategy?	✓ yes partial yes no
5. Did the review authors perform study selection in duplicate?	✓ yes ^a no
a. Nur 1 Reviewer selektierte initial anhand der Titel; danach erfolgte der weitere Screeningprozess durch 2 Personen unabhängig voneinander. Daher wurde der gesamte Prozess mit „ja“ bewertet.	

B.2 Beispiel MCL

Tabelle 22: Bewertung der Informationsbeschaffung zur systematischen Übersicht Sancho 2024 [26]

1. Did the research questions and inclusion criteria for the review include the components of PICO?	✓ yes no
4. Did the review authors use a comprehensive literature search strategy?	yes ✓ partial yes no
5. Did the review authors perform study selection in duplicate?	✓ yes no

Tabelle 23: Bewertung der Informationsbeschaffung zur systematischen Übersicht National Institute for Health and Care Excellence (NICE) 2021 [27]

1. Did the research questions and inclusion criteria for the review include the components of PICO?	✓ yes no
4. Did the review authors use a comprehensive literature search strategy?	✓ yes partial yes no
5. Did the review authors perform study selection in duplicate?	✓ yes no

Tabelle 24: Bewertung der Informationsbeschaffung zur systematischen Übersicht Wu 2024 [28]

1. Did the research questions and inclusion criteria for the review include the components of PICO?	✓ yes no
4. Did the review authors use a comprehensive literature search strategy?	✓ yes partial yes no
5. Did the review authors perform study selection in duplicate?	✓ yes no

Anhang C Studienlisten

C.1 Beispiel RRMS

C.1.1 Liste der relevanten systematischen Übersichten

1. Gonzalez-Lorenzo M, Ridley B, Minozzi S et al. Immunomodulators and immunosuppressants for relapsing-remitting multiple sclerosis: a network meta-analysis. Cochrane Database Syst Rev 2024; 1: CD011381. <https://doi.org/10.1002/14651858.CD011381.pub3>.
2. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Alemtuzumab, Cladribin, Dimethylfumarat, Fingolimod, Natalizumab, Ocrelizumab, Ofatumumab, Ozanimod, Ponesimod und Teriflunomid zur Behandlung Erwachsener mit hochaktiver schubförmig remittierender multipler Sklerose [online]. 2023 [Zugriff: 01.07.2024]. URL: https://www.iqwig.de/download/a20-60_multiple-sklerose_abschlussbericht_v1-0.pdf.
3. Landmeyer NC, Burkner PC, Wiendl H et al. Disease-modifying treatments and cognition in relapsing-remitting multiple sclerosis: A meta-analysis. Neurology 2020; 94(22): e2373-e2383. <https://doi.org/10.1212/wnl.00000000000009522>.
4. Li H, Hu F, Zhang Y et al. Comparative efficacy and acceptability of disease-modifying therapies in patients with relapsing-remitting multiple sclerosis: a systematic review and network meta-analysis. J Neurol 2020; 267(12): 3489-3498. <https://doi.org/10.1007/s00415-019-09395-w>.
5. Liu Z, Liao Q, Wen H et al. Disease modifying therapies in relapsing-remitting multiple sclerosis: A systematic review and network meta-analysis. Autoimmun Rev 2021; 20(6): 102826. <https://doi.org/10.1016/j.autrev.2021.102826>.
6. Prosperini L, Haggiag S, Ruggieri S et al. Dimethyl Fumarate or Teriflunomide for Relapsing-Remitting Multiple Sclerosis: A Meta-analysis of Post-marketing Studies. Neurotherapeutics 2023; 20(5): 1275-1283. <https://doi.org/10.1007/s13311-023-01416-x>.
7. Tramacere I, Virgili G, Perduca V et al. Adverse effects of immunotherapies for multiple sclerosis: a network meta-analysis. Cochrane Database Syst Rev 2023; 11: CD012186. <https://doi.org/10.1002/14651858.CD012186.pub2>.

C.1.2 Liste der im Volltext ausgeschlossenen systematischen Übersichten

Nicht E1

1. Hennessy B, Zierhut ML, Kracker H et al. Comparative Efficacy of Relapsing Multiple Sclerosis Therapies: Model-Based Meta-Analyses for Confirmed Disability Accumulation and Annualized Relapse Rate. Multiple Sclerosis and Related Disorders 2022; 64: 103908. <https://doi.org/10.1016/j.msard.2022.103908>.

2. Sahraian MA, Salehi AM, Jenabi E et al. Post marketing new adverse effects of oral therapies in multiple sclerosis: A systematic review. *Multiple Sclerosis and Related Disorders* 2022; 68: 104157. <https://doi.org/10.1016/j.msard.2022.104157>.

3. Valizadeh A, Fattahi MR, Sadeghi M et al. Disease-modifying therapies and T1 hypointense lesions in patients with multiple sclerosis: A systematic review and meta-analysis. *CNS Neurosci Ther* 2022; 28(5): 648-657. <https://doi.org/10.1111/cns.13815>.

Nicht E2

1. Bartosik-Psujek H, Kaczynski L, Gorecka M et al. Cladribine tablets versus other disease-modifying oral drugs in achieving no evidence of disease activity (NEDA) in multiple sclerosis- A systematic review and network meta-analysis. *Multiple Sclerosis and Related Disorders* 2021; 49: 102769. <https://doi.org/10.1016/j.msard.2021.102769>.

2. Siddiqui MK, Singh B, Attri S et al. Use of rituximab in adults with relapsing-remitting multiple sclerosis: a systematic literature review. *Curr Med Res Opin* 2020; 36(5): 809-826. <https://doi.org/10.1080/03007995.2020.1739009>.

Nicht E3

1. McGinley MP, Goldschmidt CH, Rae-Grant AD. Diagnosis and Treatment of Multiple Sclerosis: A Review. *JAMA* 2021; 325(8): 765-779. <https://doi.org/10.1001/jama.2020.26858>.

C.1.3 Liste der eingeschlossenen Primärpublikationen identifiziert durch systematische Übersichten

1. Arnold DL, Gold R, Kappos L et al. Magnetization transfer ratio in the delayed-release dimethyl fumarate DEFINE study. *J Neurol* 2014; 261(12): 2429-2437. <https://doi.org/10.1007/s00415-014-7504-7>.

2. Arnold DL, Gold R, Kappos L et al. Effects of delayed-release dimethyl fumarate on MRI measures in the Phase 3 DEFINE study. *J Neurol* 2014; 261(9): 1794-1802. <https://doi.org/10.1007/s00415-014-7412-x>.

3. Bar-Or A, Gold R, Kappos L et al. Clinical efficacy of BG-12 (dimethyl fumarate) in patients with relapsing-remitting multiple sclerosis: subgroup analyses of the DEFINE study. *J Neurol* 2013; 260(9): 2297-2305. <https://doi.org/10.1007/s00415-013-6954-7>.

4. Braune S, Grimm S, van Hovell P et al. Comparative effectiveness of delayed-release dimethyl fumarate versus interferon, glatiramer acetate, teriflunomide, or fingolimod: results from the German NeuroTransData registry. *J Neurol* 2018; 265(12): 2980-2992. <https://doi.org/10.1007/s00415-018-9083-5>.

5. Buron MD, Chalmer TA, Sellebjerg F et al. Comparative effectiveness of teriflunomide and dimethyl fumarate: A nationwide cohort study. *Neurology* 2019; 92(16): e1811-e1820. <https://doi.org/10.1212/wnl.00000000000007314>.

6. Conde S, Moisset X, Pereira B et al. Dimethyl fumarate and teriflunomide for multiple sclerosis in a real-life setting: a French retrospective cohort study. *Eur J Neurol* 2019; 26(3): 460-467. <https://doi.org/10.1111/ene.13839>.
7. Cree BAC, Goldman MD, Corboy JR et al. Efficacy and Safety of 2 Fingolimod Doses vs Glatiramer Acetate for the Treatment of Patients With Relapsing-Remitting Multiple Sclerosis; a Randomized Clinical Trial. *JAMA neurology* 2020; 78(1): 1-13. <https://doi.org/10.1001/jamaneurol.2020.2950>.
8. D'Amico E, Zanghi A, Sciandra M et al. Dimethyl fumarate vs Teriflunomide: an Italian time-to-event data analysis. *J Neurol* 2020; 267(10): 3008-3020. <https://doi.org/10.1007/s00415-020-09959-1>.
9. Davis MD, Ashtamker N, Steinerman JR et al. Time course of glatiramer acetate efficacy in patients with RRMS in the GALA study. *Neurol Neuroimmunol Neuroinflamm* 2017; 4(2): e327. <https://doi.org/10.1212/NXI.0000000000000327>.
10. Fernández Ó, Giovannoni G, Fox RJ et al. Efficacy and Safety of Delayed-release Dimethyl Fumarate for Relapsing-remitting Multiple Sclerosis in Prior Interferon Users; an Integrated Analysis of DEFINE and CONFIRM. *Clin Ther* 2017; 39(8): 1671-1679. <https://doi.org/10.1016/j.clinthera.2017.06.012>.
11. 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>.
12. Gold R, Kappos L, Arnold DL et al. Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis. *N Engl J Med* 2012; 367(12): 1098-1107. <https://doi.org/10.1056/NEJMoa1114287>.
13. Guger M, Enzinger C, Leutmezer F et al. Oral therapies for treatment of relapsing-remitting multiple sclerosis in Austria: a 2-year comparison using an inverse probability weighting method. *J Neurol* 2020; 267(7): 2090-2100. <https://doi.org/10.1007/s00415-020-09811-6>.
14. Havrdova E, Giovannoni G, Gold R et al. Effect of delayed-release dimethyl fumarate on no evidence of disease activity in relapsing-remitting multiple sclerosis; integrated analysis of the phase III DEFINE and CONFIRM studies. *Eur J Neurol* 2017; 24(5): 726-733. <https://doi.org/10.1111/ene.13272>.
15. Hillert J, Tsai JA, Nouhi M et al. A comparative study of teriflunomide and dimethyl fumarate within the Swedish MS Registry. *Multiple Sclerosis Journal* 2022; 28(2): 237-246. <https://doi.org/10.1177/13524585211019649>.

16. Hutchinson M, Fox RJ, Miller DH et al. Clinical efficacy of BG-12 (dimethyl fumarate) in patients with relapsing-remitting multiple sclerosis; subgroup analyses of the CONFIRM study. *J Neurol* 2013; 260(9): 2286-2296. <https://doi.org/10.1007/s00415-013-6968-1>.
17. Kalincik T, Kubala Havrdova E, Horakova D et al. Comparison of fingolimod, dimethyl fumarate and teriflunomide for multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2019; 90(4): 458-468. <https://doi.org/10.1136/jnnp-2018-319831>.
18. Kappos L, Gold R, Arnold DL et al. Quality of life outcomes with BG-12 (dimethyl fumarate) in patients with relapsing-remitting multiple sclerosis; the DEFINE study. *Multiple Sclerosis Journal* 2014; 20(2): 243-252. <https://doi.org/10.1177/1352458513507817>.
19. Khan O, Rieckmann P, Boyko A et al. Three times weekly glatiramer acetate in relapsing-remitting multiple sclerosis. *Ann Neurol* 2013; 73(6): 705-713. <https://doi.org/10.1002/ana.23938>.
20. Kita M, Fox RJ, Phillips JT et al. Effects of BG-12 (dimethyl fumarate) on health-related quality of life in patients with relapsing-remitting multiple sclerosis; findings from the CONFIRM study. *Multiple Sclerosis Journal* 2014; 20(2): 253-257. <https://doi.org/10.1177/1352458513507818>.
21. Laplaud DA, Casey R, Barbin L et al. Comparative effectiveness of teriflunomide vs dimethyl fumarate in multiple sclerosis. *Neurology* 2019; 93(7): e635-e646. <https://doi.org/10.1212/wnl.00000000000007938>.
22. Mikol DD, Barkhof F, Chang P et al. Comparison of subcutaneous interferon beta-1a with glatiramer acetate in patients with relapsing multiple sclerosis (the REbif vs Glatiramer Acetate in Relapsing MS Disease [REGARD] study); a multicentre, randomised, parallel, open-label trial. *The Lancet Neurology* 2008; 7(10): 903-914. [https://doi.org/10.1016/S1474-4422\(08\)70200-X](https://doi.org/10.1016/S1474-4422(08)70200-X).
23. Miller DH, Fox RJ, Phillips JT et al. Effects of delayed-release dimethyl fumarate on MRI measures in the phase 3 CONFIRM study. *Neurology* 2015; 84(11): 1145-1152. <https://doi.org/10.1212/WNL.0000000000001360>.
24. Nehzat N, Mirmosayyeb O, Barzegar M et al. Comparable Efficacy and Safety of Teriflunomide versus Dimethyl Fumarate for the Treatment of Relapsing-Remitting Multiple Sclerosis. *Neurol Res Int* 2021; 2021: 6679197. <https://doi.org/10.1155/2021/6679197>.
25. O'Connor P, Filippi M, Arnason B et al. 250 microg or 500 microg interferon beta-1b versus 20 mg glatiramer acetate in relapsing-remitting multiple sclerosis; a prospective, randomised, multicentre study. *Lancet Neurol* 2009; 8(10): 889-897. [https://doi.org/10.1016/S1474-4422\(09\)70226-1](https://doi.org/10.1016/S1474-4422(09)70226-1).

26. Prosperini L, Cortese A, Lucchini M et al. Exit strategies for "needle fatigue" in multiple sclerosis: a propensity score-matched comparison study. J Neurol 2020; 267(3): 694-702. <https://doi.org/10.1007/s00415-019-09625-1>.
27. Schippling S, O'Connor P, Knappertz V et al. Incidence and course of depression in multiple sclerosis in the multinational BEYOND trial. J Neurol 2016; 263(7): 1418-1426. <https://doi.org/10.1007/s00415-016-8146-8>.
28. Vermersch P, Suchet L, Colamarino R et al. An analysis of first-line disease-modifying therapies in patients with relapsing-remitting multiple sclerosis using the French nationwide health claims database from 2014-2017. Mult Scler Relat Disord 2020; 46: 102521. <https://doi.org/10.1016/j.msard.2020.102521>.

C.1.4 Liste der im Volltext ausgeschlossenen Primärpublikationen identifiziert durch systematische Übersichten

Nicht E1

1. Boster A, Nicholas J, Wu N et al. Comparative Effectiveness Research of Disease-Modifying Therapies for the Management of Multiple Sclerosis: Analysis of a Large Health Insurance Claims Database. Neurology and Therapy 2017; 6(1): 91-102. <https://doi.org/10.1007/s40120-017-0064-x>.
2. Brown RA, Narayanan S, Stikov N et al. MTR recovery in brain lesions in the BECOME study of glatiramer acetate vs interferon β -1b. Neurology 2016; 87(9): 905-911. <https://doi.org/10.1212/WNL.0000000000003043>.
3. Cadavid D, Wolansky LJ, Skurnick J et al. Efficacy of treatment of MS with IFN β -1b or glatiramer acetate by monthly brain MRI in the BECOME study. Neurology 2009; 72(23): 1976-1983. <https://doi.org/10.1212/01.wnl.0000345970.73354.17>.
4. Cheriyan J, Kim S, Wolansky LJ et al. Impact of inflammation on brain volume in multiple sclerosis. Arch Neurol 2012; 69(1): 82-88. <https://doi.org/10.1001/archneurol.2011.674>.
5. Ontaneda D, Nicholas J, Carraro M et al. Comparative effectiveness of dimethyl fumarate versus fingolimod and teriflunomide among MS patients switching from first-generation platform therapies in the US. Multiple Sclerosis and Related Disorders 2019; 27: 101-111. <https://doi.org/10.1016/j.msard.2018.09.038>.
6. Zivadinov R, Kresa-Reahl K, Weinstock-Guttman B et al. Comparative effectiveness of teriflunomide and dimethyl fumarate in patients with relapsing forms of MS in the retrospective real-world Teri-RADAR study. J Comp Eff Res 2019; 8(5): 305-316. <https://doi.org/10.2217/ce-2018-0135>.

Nicht E4

1. Zivadinov R, Dwyer M, Barkay H et al. Effect of glatiramer acetate three-times weekly on the evolution of new, active multiple sclerosis lesions into T1-hypointense "black holes"; a post hoc magnetic resonance imaging analysis. J Neurol 2015; 262(3): 648-653.

<https://doi.org/10.1007/s00415-014-7616-0>.

2. Zivadinov R, Dwyer MG, Ramasamy DP et al. The Effect of Three Times a Week Glatiramer Acetate on Cerebral T1 Hypointense Lesions in Relapsing-Remitting Multiple Sclerosis. J Neuroimaging 2015; 25(6): 989-995. <https://doi.org/10.1111/jon.12293>.

C.1.5 Liste der eingeschlossenen Primärpublikationen aus dem Datensatz der externen Sachverständigen

1. Alsop J, Medin J, Cornelissen C et al. 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. PLoS One 2017; 12(5):

e0173353. <https://doi.org/10.1371/journal.pone.0173353>.

2. Bergvall N, Makin C, Lahoz R et al. Relapse rates in patients with multiple sclerosis switching from interferon to fingolimod or glatiramer acetate: a US claims database study. PLoS One 2014; 9(2): e88472. <https://doi.org/10.1371/journal.pone.0088472>.

3. Boggild M, Palace J, Barton P et al. Multiple sclerosis risk sharing scheme: two year results of clinical cohort study with historical comparator. BMJ 2009; 339: b4677.

<https://doi.org/10.1136/bmj.b4677>.

4. Boster A, Nicholas J, Wu N et al. Comparative Effectiveness Research of Disease-Modifying Therapies for the Management of Multiple Sclerosis: Analysis of a Large Health Insurance Claims Database. Neurol Ther 2017; 6(1): 91-102. <https://doi.org/10.1007/s40120-017-0064-x>.

5. Bovis F, Kalincik T, Lublin F et al. Treatment Response Score to Glatiramer Acetate or Interferon Beta-1a. Neurology 2021; 96(2): e214-e227.

<https://doi.org/10.1212/WNL.00000000000010991>.

6. Buron MD, Kalincik T, Sellebjerg F et al. Effect of lateral therapy switches to oral moderate-efficacy drugs in multiple sclerosis: a nationwide cohort study. J Neurol Neurosurg Psychiatry 2021; 92(5): 556-562. <https://doi.org/10.1136/jnnp-2020-324869>.

7. Ciplea AI, Kurzeja A, Thiel S et al. Safety evaluations of offspring breastfed by mothers receiving glatiramer acetate for relapsing multiple sclerosis. Mult Scler Relat Disord 2023; 75: 104771. <https://doi.org/10.1016/j.msard.2023.104771>.

8. Diouf I, Malpas CB, Sharmin S et al. Effectiveness of multiple disease-modifying therapies in relapsing-remitting multiple sclerosis: causal inference to emulate a multiarm randomised trial. *J Neurol Neurosurg Psychiatry* 2023; 94(12): 1004-1011. <https://doi.org/10.1136/jnnp-2023-331499>.
9. Ferraro D, Camera V, Baldi E et al. First-line disease-modifying drugs in relapsing-remitting multiple sclerosis: an Italian real-life multicenter study on persistence. *Curr Med Res Opin* 2018; 34(10): 1803-1807. <https://doi.org/10.1080/03007995.2018.1451311>.
10. Granqvist M, Burman J, Gunnarsson M et al. Comparative effectiveness of dimethyl fumarate as the initial and secondary treatment for MS. *Mult Scler* 2020; 26(12): 1532-1539. <https://doi.org/10.1177/1352458519866600>.
11. Haas J, Firzlaff M. Twenty-four-month comparison of immunomodulatory treatments - a retrospective open label study in 308 RRMS patients treated with beta interferons or glatiramer acetate (Copaxone). *Eur J Neurol* 2005; 12(6): 425-431. <https://doi.org/10.1111/j.1468-1331.2005.00936.x>.
12. Hellwig K, Thiel S, Haben S et al. Glatiramer acetate or IFN-beta bridging therapy in women with relapsing multiple sclerosis planning a pregnancy. *Neurodegener Dis Manag* 2023; 13(4): 223-234. <https://doi.org/10.2217/nmt-2022-0032>.
13. Iaffaldano P, Lucisano G, Pozzilli C et al. Fingolimod versus interferon beta/glatiramer acetate after natalizumab suspension in multiple sclerosis. *Brain* 2015; 138(Pt 11): 3275-3286. <https://doi.org/10.1093/brain/awv260>.
14. Izquierdo G, Garcia-Agua Soler N, Rus M et al. Effectiveness of glatiramer acetate compared to other multiple sclerosis therapies. *Brain Behav* 2015; 5(6): e00337. <https://doi.org/10.1002/brb3.337>.
15. Mesidor M, Sylvestre MP, Rousseau MC et al. 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. <https://doi.org/10.1016/j.msard.2021.103050>.
16. Mitsikostas D, Bakirtzis C, Nikolaidis I et al. Quality of life in people with multiple sclerosis receiving glatiramer acetate or interferon in Greek clinical practice. *Neurodegener Dis Manag* 2022; 12(6): 311-322. <https://doi.org/10.2217/nmt-2022-0004>.
17. Palace J, Duddy M, Lawton 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(3): 251-260. <https://doi.org/10.1136/jnnp-2018-318360>.

18. Pfeuffer S, Kerschke L, Ruck T et al. Teriflunomide treatment is associated with optic nerve recovery in early multiple sclerosis. *Ther Adv Neurol Disord* 2021; 14: 1756286421997372. <https://doi.org/10.1177/1756286421997372>.
19. Reder AT, Arndt N, Roman C et al. 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. *Mult Scler Relat Disord* 2021; 51: 102935. <https://doi.org/10.1016/j.msard.2021.102935>.
20. Rieckmann P, Zivadinov R, Boyko A et al. 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. *Mult Scler J Exp Transl Clin* 2021; 7(4): 20552173211061550. <https://doi.org/10.1177/20552173211061550>.
21. Rojas JI, Patrucco L, Alonso R et al. Effectiveness and Safety of Early High-Efficacy Versus Escalation Therapy in Relapsing-Remitting Multiple Sclerosis in Argentina. *Clin Neuropharmacol* 2022; 45(3): 45-51. <https://doi.org/10.1097/WNF.0000000000000503>.
22. Romeo MAL, Martinelli V, Dalla Costa G et al. Assessing the role of innovative therapeutic paradigm on multiple sclerosis treatment response. *Acta Neurol Scand* 2018; 138(5): 447-453. <https://doi.org/10.1111/ane.12999>.
23. Rossi S, Motta C, Studer V et al. Treatment options to reduce disease activity after natalizumab: paradoxical effects of corticosteroids. *CNS Neurosci Ther* 2014; 20(8): 748-753. <https://doi.org/10.1111/cns.12282>.
24. Rovaris M, Comi G, Rocca MA et al. Short-term brain volume change in relapsing-remitting multiple sclerosis: effect of glatiramer acetate and implications. *Brain* 2001; 124(Pt 9): 1803-1812. <https://doi.org/10.1093/brain/124.9.1803>.
25. Scott TF, Su R, Xiong K et al. Matching comparisons of therapeutic efficacy suggest better clinical outcomes for patients treated with peginterferon beta-1a than with glatiramer acetate. *Ther Adv Neurol Disord* 2021; 14: 1756286420975916. <https://doi.org/10.1177/1756286420975916>.
26. Signori A, Sacca F, Lanzillo R et al. Cladribine vs other drugs in MS: Merging randomized trial with real-life data. *Neurol Neuroimmunol Neuroinflamm* 2020; 7(6). <https://doi.org/10.1212/NXI.0000000000000878>.
27. Spelman T, Kalincik T, Jokubaitis V et al. Comparative efficacy of first-line natalizumab vs IFN-beta or glatiramer acetate in relapsing MS. *Neurol Clin Pract* 2016; 6(2): 102-115. <https://doi.org/10.1212/CPJ.0000000000000227>.

28. von Wyl V, Benkert P, Moser A et al. Disability progression in relapse-free multiple sclerosis patients on fingolimod versus interferon-beta/glatiramer acetate. *Mult Scler* 2021; 27(3): 439-448. <https://doi.org/10.1177/1352458520918489>.

29. Yeh WZ, Widyastuti PA, Van der Walt A et al. Natalizumab, Fingolimod and Dimethyl Fumarate Use and Pregnancy-Related Relapse and Disability in Women With Multiple Sclerosis. *Neurology* 2021; 96(24): e2989-e3002. <https://doi.org/10.1212/WNL.00000000000012084>.

C.2 Beispiel MCL

C.2.1 Liste der relevanten systematischen Übersichten

1. Monga N, Garside J, Davids MS et al. Systematic Literature Review of Economic Evaluations, Costs/Resource Use, and Quality of Life in Patients with Mantle Cell Lymphoma. *PharmacoEconomics Open* 2021; 5(2): 175-186. <https://doi.org/10.1007/s41669-020-00231-w>.

2. National Institute for Health and Care Excellence. Brexucabtagene autoleucel for treating relapsed or refractory mantle cell lymphoma; technology appraisal guidance; TA677 [online]. 2021 [Zugriff: 21.08.2024]. URL: <https://www.nice.org.uk/guidance/ta677>.

3. Roufarshbaf M, Javeri M, Akbari V et al. Efficacy and safety of ibrutinib in mantle cell lymphoma: A systematic review and meta-analysis. *Daru: Journal of Faculty of Pharmacy, Tehran University of Medical Sciences* 2022; 30(2): 367-378. <https://doi.org/10.1007/s40199-022-00444-w>.

4. Sancho JM, Sorigue M, Rubio-Azpeitia E. Real-World Evidence of Relapsed/Refractory Mantle Cell Lymphoma Patients and Treatments: A Systematic Review. *J Blood Med* 2024; 15: 239-254. <https://doi.org/10.2147/jbm.S463946>.

5. Wu JJ, Wade SW, Itani T et al. Unmet needs in relapsed/refractory mantle cell lymphoma (r/r MCL) post-covalent Bruton tyrosine kinase inhibitor (BTKi): a systematic literature review and meta-analysis. *Leuk Lymphoma* 2024: 1-14. <https://doi.org/10.1080/10428194.2024.2369653>.

C.2.2 Liste der im Volltext ausgeschlossenen systematischen Übersichten

Nicht E1

1. Alnassfan T, Cox-Pridmore MJ, Taktak A et al. Mantle cell lymphoma treatment options for elderly/unfit patients: A systematic review. *EJHaem* 2022; 3(1): 276-290. <https://doi.org/10.1002/jha2.311>.

2. Jing C, Zhao A, Wang J et al. Efficacy of front-line immunochemotherapy for transplant-ineligible mantle cell lymphoma: A network meta-analysis of randomized controlled trials. *Cancer Medicine* 2023; 12(14): 15107-15116. <https://doi.org/10.1002/cam4.6183>.

3. Li SJ, Hao J, Mao Y et al. Effects of the Proteasome Inhibitor Bortezomib in Combination with Chemotherapy for the Treatment of Mantle Cell Lymphoma: A Meta-analysis. Turkish Journal of Haematology 2020; 37(1): 13-19.

<https://doi.org/10.4274/tjh.galenos.2019.2019.0128>.

4. Liu H, Shi X, Fang H et al. First-Line Autologous Stem Cell Transplantation for Mantle Cell Lymphoma: A Systematic Analysis and Treatment Recommendation. Front Oncol 2022; 12: 881346. <https://doi.org/10.3389/fonc.2022.881346>.

Nicht E2

1. Monga N, Garside J, Quigley J et al. Systematic literature review of the global burden of illness of mantle cell lymphoma. Curr Med Res Opin 2020; 36(5): 843-852.

<https://doi.org/10.1080/03007995.2020.1742101>.

Nicht E3

1. Agency for Care Effectiveness. Acalabrutinib, bortezomib and ibrutinib for treating mantle cell lymphoma [online]. 2022 [Zugriff: 21.08.2024]. URL: <https://www.ace-hta.gov.sg/healthcare-professionals/ace-technology-guidances/details/acalabrutinib-bortezomib-and-ibrutinib-for-treating-mantle-cell-lymphoma>.

2. Anderson MK, Torosyan A, Halford Z. Brexucabtagene Autoleucel: A Novel Chimeric Antigen Receptor T-cell Therapy for the Treatment of Mantle Cell Lymphoma. Ann Pharmacother 2022; 56(5): 609-619. <https://doi.org/10.1177/10600280211026338>.

3. Bhella S, Varela NP, Aw A et al. First-line therapy, autologous stem-cell transplantation, and post-transplantation maintenance in the management of newly diagnosed mantle cell lymphoma. Curr Oncol 2020; 27(6): e632-e644. <https://doi.org/10.3747/co.27.7053>.

4. Buege MJ, Kumar A, Dixon BN et al. Management of Mantle Cell Lymphoma in the Era of Novel Oral Agents. Ann Pharmacother 2020; 54(9): 879-898. <https://doi.org/10.1177/1060028020909117>.

5. Edina BC, Rinaldi I. Effectiveness of Bendamustine-Rituximab Compared to R-CHOP/R-CVP as a First-Line Treatment of Indolent Non-Hodgkin's Lymphoma or Mantle-Cell Lymphoma. Acta Med Indones 2022; 54(2): 316-323.

6. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Autologe Anti-CD19-transduzierte CD3-positive Zellen (Mantelzell-Lymphom) [online]. 2021. URL: https://www.iqwig.de/download/g21-08_autologe-anti-cd19-transduzierte-cd3-positive-zellen_bewertung-35a-absatz-1-satz-11-sgb-v_v1-0.pdf.

7. Sheng Z, Wang L. Superiority of ibrutinib plus bendamustine and rituximab in newly diagnosed patients with mantle-cell lymphoma ineligible for intensive therapy: A network meta-analysis. Eur J Haematol 2023; 110(6): 602-607. <https://doi.org/10.1111/ejh.13953>.

C.2.3 Liste der eingeschlossenen Primärpublikationen identifiziert durch systematische Übersichten

1. Broccoli A, Casadei B, Morigi A et al. Italian real life experience with ibrutinib; results of a large observational study on 77 relapsed/refractory mantle cell lymphoma. *Oncotarget* 2018; 9(34): 23443-23450. <https://doi.org/10.18632/oncotarget.25215>.
2. Cencini E, Mecacci B, Morelli F et al. Ibrutinib in patients with relapsed/refractory mantle cell lymphoma: a real-life, retrospective, multicenter trial on behalf of the "RTL" (regional Tuscan lymphoma network). *Am J Blood Res* 2021; 11(4): 373-383.
3. Cheah CY, Chihara D, Romaguera JE et al. Patients with mantle cell lymphoma failing ibrutinib are unlikely to respond to salvage chemotherapy and have poor outcomes. *Ann Oncol* 2015; 26(6): 1175-1179. <https://doi.org/10.1093/annonc/mdv111>.
4. Dreyling M, Jurczak W, Jerkeman M et al. Ibrutinib versus temsirolimus in patients with relapsed or refractory mantle-cell lymphoma; an international, randomised, open-label, phase 3 study. *Lancet* 2016; 387(10020): 770-778. [https://doi.org/10.1016/s0140-6736\(15\)00667-4](https://doi.org/10.1016/s0140-6736(15)00667-4).
5. Epperla N, Hamadani M, Cashen AF et al. Predictive factors and outcomes for ibrutinib therapy in relapsed/refractory mantle cell lymphoma—a “real world” study. *Hematol Oncol* 2017; 35(4): 528-535. <https://doi.org/10.1002/hon.2380>.
6. Eyre TA, Walter HS, Iyengar S et al. Efficacy of venetoclax monotherapy in patients with relapsed, refractory mantle cell lymphoma after Bruton tyrosine kinase inhibitor therapy. *Haematologica* 2019; 104(2): e68-e71. <https://doi.org/10.3324/haematol.2018.198812>.
7. García-Noblejas A, Martínez Chamorro C, Navarro Matilla B et al. Bendamustine as salvage treatment for patients with relapsed or refractory mantle cell lymphoma patients; a retrospective study of the Spanish experience. *Ann Hematol* 2014; 93(9): 1551-1558. <https://doi.org/10.1007/s00277-014-2077-1>.
8. Hess G, Dreyling M, Oberic L et al. Real-world experience among patients with relapsed/refractory mantle cell lymphoma after Bruton tyrosine kinase inhibitor failure in Europe; The SCHOLAR-2 retrospective chart review study. *Br J Haematol* 2023; 202(4): 749-759. <https://doi.org/10.1111/bjh.18519>.
9. Iacoboni G, Rejeski K, Villacampa G et al. Real-world evidence of brexucabtagene autoleucel for the treatment of relapsed or refractory mantle cell lymphoma. *Blood advances* 2022; 6(12): 3606-3610. <https://doi.org/10.1182/bloodadvances.2021006922>.
10. Jain P, Kanagal-Shamanna R, Zhang S et al. Long-term outcomes and mutation profiling of patients with mantle cell lymphoma (MCL) who discontinued ibrutinib. *Br J Haematol* 2018; 183(4): 578-587. <https://doi.org/10.1111/bjh.15567>.

11. Jeon Y-W, Yoon S, Min GJ et al. Clinical outcomes for ibrutinib in relapsed or refractory mantle cell lymphoma in real-world experience. *Cancer medicine* 2019; 8(16): 6860-6870. <https://doi.org/10.1002/cam4.2565>.
12. Martin P, Maddocks K, Leonard JP et al. Postibrutinib outcomes in patients with mantle cell lymphoma. *Blood* 2016; 127(12): 1559-1563. <https://doi.org/10.1182/blood-2015-10-673145>.
13. McCulloch R, Lewis D, Crosbie N et al. Ibrutinib for mantle cell lymphoma at first relapse; a United Kingdom real-world analysis of outcomes in 211 patients. *Br J Haematol* 2021; 193(2): 290-298. <https://doi.org/10.1111/bjh.17363>.
14. McCulloch R, Visco C, Eyre TA et al. Efficacy of R-BAC in relapsed, refractory mantle cell lymphoma post BTK inhibitor therapy. *Br J Haematol* 2020; 189(4): 684-688. <https://doi.org/10.1111/bjh.16416>.
15. National Institute for Health and Care Excellence. Single technology appraisal; Autologous anti-CD19-transduced CD3+ cells for treating relapsed or refractory mantle cell lymphoma [ID1313] [online]. 2021 [Zugriff: 09.07.2026]. URL: <https://www.nice.org.uk/guidance/ta677>.
16. Rai S, Tanizawa Y, Cai Z et al. Outcomes for Recurrent Mantle Cell Lymphoma Post-Ibrutinib Therapy; A Retrospective Cohort Study from a Japanese Administrative Database. *Adv Ther* 2022; 39(10): 4792-4807. <https://doi.org/10.1007/s12325-022-02258-3>.
17. Sancho J-M, Marín-Niebla A, Fernández S et al. IBRORS-MCL study; a Spanish retrospective and observational study of relapsed/refractory mantle-cell lymphoma treated with ibrutinib in routine clinical practice. *Int J Hematol* 2022; 116(3): 381-392. <https://doi.org/10.1007/s12185-022-03367-z>.
18. Sharman J, Kabadi SM, Clark J et al. Treatment patterns and outcomes among mantle cell lymphoma patients treated with ibrutinib in the United States; a retrospective electronic medical record database and chart review study. *Br J Haematol* 2020; 192(4): 737-746. <https://doi.org/10.1111/bjh.16922>.
19. Tucker D, Morley N, MacLean P et al. The 5-year follow-up of a real-world observational study of patients in the United Kingdom and Ireland receiving ibrutinib for relapsed/refractory mantle cell lymphoma. *Br J Haematol* 2020; 192(6): 1035-1038. <https://doi.org/10.1111/bjh.16739>.
20. Visco C, Di Rocco A, Evangelista A et al. Outcomes in first relapsed-refractory younger patients with mantle cell lymphoma; results from the MANTLE-FIRST study. *Leukemia* 2020; 35(3): 787-795. <https://doi.org/10.1038/s41375-020-01013-3>.

21. Wang M, Munoz J, Goy A et al. Three-Year Follow-Up of KTE-X19 in Patients With Relapsed/Refractory Mantle Cell Lymphoma, Including High-Risk Subgroups, in the ZUMA-2 Study. J Clin Oncol 2023; 41(3): 555-567. <https://doi.org/10.1200/JCO.21.02370>.
22. Wang M, Munoz J, Goy A et al. KTE-X19 CAR T-Cell Therapy in Relapsed or Refractory Mantle-Cell Lymphoma. N Engl J Med 2020; 382(14): 1331-1342. <https://doi.org/10.1056/NEJMoa1914347>.
23. Wang Y, Jain P, Locke FL et al. Brexucabtagene Autoleucel for Relapsed or Refractory Mantle Cell Lymphoma in Standard-of-Care Practice; Results From the US Lymphoma CAR T Consortium. J Clin Oncol 2023; 41(14): 2594-2606. <https://doi.org/10.1200/JCO.22.01797>.
24. Yi JH, Kim SJ, Yoon DH et al. Real-world outcomes of ibrutinib therapy in Korean patients with relapsed or refractory mantle cell lymphoma; a multicenter, retrospective analysis. Cancer Commun 2021; 41(3): 275-278. <https://doi.org/10.1002/cac2.12150>.

C.2.4 Liste der im Volltext ausgeschlossenen Primärpublikationen identifiziert durch systematische Übersichten

Nicht E1

1. Ruan J, Martin P, Shah B et al. Lenalidomide plus Rituximab as Initial Treatment for Mantle-Cell Lymphoma. N Engl J Med 2015; 373(19): 1835-1844. <https://doi.org/10.1056/NEJMoa1505237>.
2. Smith A, Roman E, Appleton S et al. Impact of novel therapies for mantle cell lymphoma in the real world setting; a report from the UK's Haematological Malignancy Research Network (HMRN). Br J Haematol 2018; 181(2): 215-228. <https://doi.org/10.1111/bjh.15170>.
3. Warsch S, Hosein PJ, Maeda LS et al. A retrospective study evaluating the efficacy and safety of bendamustine in the treatment of mantle cell lymphoma. Leuk Lymphoma 2012; 53(7): 1299-1305. <https://doi.org/10.3109/10428194.2011.649476>.
4. Widmer F, Balabanov S, Soldini D et al. R-hyper-CVAD versus R-CHOP/cytarabine with high-dose therapy and autologous haematopoietic stem cell support in fit patients with mantle cell lymphoma; 20 years of single-center experience. Ann Hematol 2017; 97(2): 277-287. <https://doi.org/10.1007/s00277-017-3180-x>.
5. Zinzani PL, Rigacci L, Cox MC et al. Lenalidomide monotherapy in heavily pretreated patients with non-Hodgkin lymphoma; an Italian observational multicenter retrospective study in daily clinical practice. Leuk Lymphoma 2014; 56(6): 1671-1676. <https://doi.org/10.3109/10428194.2014.964702>.

Nicht E2

1. Regny C, Oberic L, Guillaume M et al. Clinical efficacy of the RiBVD regimen for refractory/relapsed (r/r) mantle cell lymphoma (MCL) patients: A retrospective study of the LYSA group. Hemasphere 2019; 3(Suppl 1): 193. <https://doi.org/10.1002/j.2572-9241.2019.tb00081.x>.

Nicht E5

1. Skarbnik AP, Ma E, Lafeuille M-H et al. Treatment patterns and outcomes with subcutaneous bortezomib in patients with relapsed mantle cell lymphoma; a retrospective, observational study of patient medical records from US community oncology practices. Leuk Lymphoma 2017; 58(8): 1968-1972. <https://doi.org/10.1080/10428194.2016.1272688>.

2. Stefoni V, Pellegrini C, Broccoli A et al. Lenalidomide in Pretreated Mantle Cell Lymphoma Patients; An Italian Observational Multicenter Retrospective Study in Daily Clinical Practice (the Lenamant Study). Oncologist 2018; 23(9): 1033-1038. <https://doi.org/10.1634/theoncologist.2017-0597>.

3. Wang M, Schuster SJ, Phillips T et al. Observational study of lenalidomide in patients with mantle cell lymphoma who relapsed/progressed after or were refractory/intolerant to ibrutinib (MCL-004). J Hematol Oncol 2017; 10(1): 171-171. <https://doi.org/10.1186/s13045-017-0537-5>.

4. Zhang Y, Liu P, Cai J et al. Ibrutinib as monotherapy versus combination therapy in Chinese patients with relapsed/refractory mantle cell lymphoma; A multicenter study. Cancer medicine 2022; 11(22): 4134-4145. <https://doi.org/10.1002/cam4.4765>.

Nicht E7

1. McCulloch R, Visco C, Frewin R et al. Efficacy of R-BAC Immunochemotherapy in Patients with Relapsed, Refractory Mantle cell Lymphoma Post BTK Inhibitor Therapy. Hemashpere 2019; 3(Suppl 1): 573. <https://doi.org/10.1097/01.HS9.0000563300.27415.88>.

Anhang D Anwendung eines Sättigungskriteriums

Wie in Abschnitt 5.1.2 beschrieben, wurde neben der Simulation einer zufälligen Anordnung der Publikationen geprüft, zu welchen Ergebnissen eine inhaltlich sinnvolle Reihenfolge führt. In Tabelle 25 sind ergänzend die Ergebnisse einer Anordnung der Publikationen nach Publikationsdatum (aufsteigend und absteigend) dargestellt. Auffällig ist, dass bei der Anordnung von älteren zu aktuelleren Publikationen auch mit einem Sättigungskriterium von 5 alle 125 Confounder identifiziert werden. Bei der Reihenfolge von aktuellen zu älteren Publikationen zeigt sich ein ähnliches Ergebnis wie in der Simulation einer zufälligen Reihenfolge, nämlich einem relevanten Informationsverlust bei Anwendung eines Sättigungskriteriums von 5. Insgesamt stellen die Ergebnisse dieses einzelnen Beispiels die in Abschnitt 5.1.2 getroffenen Schlussfolgerungen nicht infrage.

Tabelle 25: Sättigungskriterium – Ergebnis der zeitlichen Anordnung der 80 Publikationen bei Anwendung eines Sättigungskriteriums von 15, 10 oder 5

Sättigungskriterium	Anzahl Studien	Anzahl Confounder	Verlust N (%) ^a
Reihenfolge basierend auf dem Veröffentlichungsjahr: von aktuelleren zu älteren Publikationen			
15	80	125	0 (0)
10	80	125	0 (0)
5	52	103	22 (18)
Reihenfolge basierend auf dem Veröffentlichungsjahr: von älteren zu aktuelleren Publikationen			
15	80	125	0 (0)
10	80	125	0 (0)
5	80	125	0 (0)
a. Anzahl an Confounder-Variablen, welche nicht identifiziert werden.			

Anhang E Erarbeitung eines gerichteten azyklischen Graphen

E.1 Zeitliches Raster

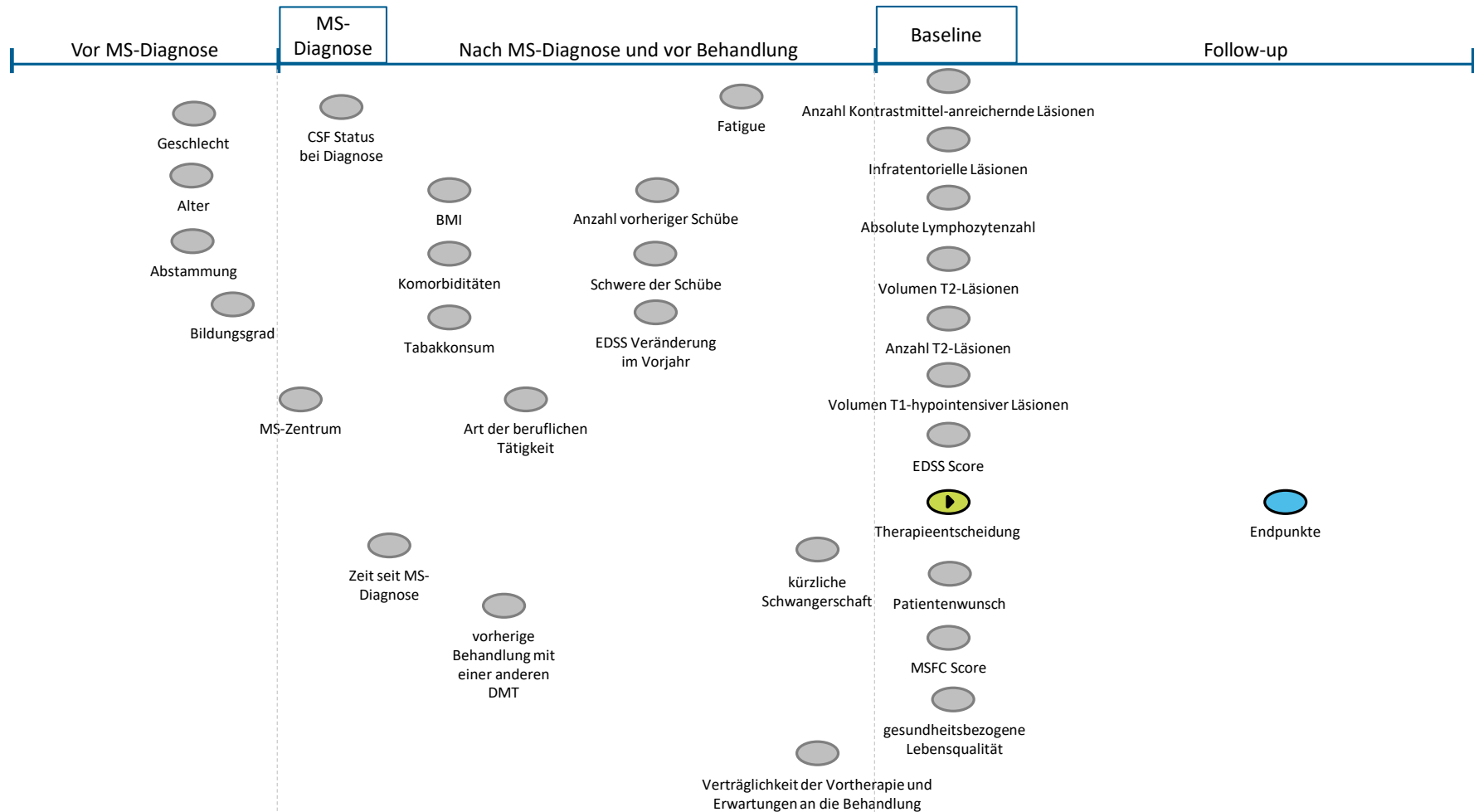


Abbildung 10: Anordnung der Confounder-Variablen im zeitlichen Raster (Ausgangspunkt für den DAG-Workshop)

E.2 Matrix der Zusammenhänge und DAGitty Model Code

E.2.1 Matrix zur Dokumentation der Zusammenhänge zwischen den Variablen

	Therapieentscheidung	Endpunkt	Geschlecht	Abstammung	Alter	Bildungsgrad	Art der beruflichen Tätigkeit	Patientenwunsch	Tabakkonsum	BMI	Fatigue	Komorbiditäten	vorherige Behandlung mit einer anderen DMT	Zeit seit der Erstmanifestation	Zeit seit der MS-Diagnose	Supernode Veränderung Krankheitsaktivität im Vorjahr	Supernode Krankheitsaktivität Baseline	Verträglichkeit der Vortherapie	Erwartungen an die Behandlung	kürzliche Schwangerschaft	absolute Lymphozytenzahl	CSF Status	MS-Zentrum	gesundheitsbezogene Lebensqualität
Therapieentscheidung																								
Endpunkt																								
Geschlecht																								
Abstammung																								
Alter																								
Bildungsgrad																								
Art der beruflichen Tätigkeit																								
Patientenwunsch																								
Tabakkonsum																								
BMI																								
Fatigue																								
Komorbiditäten																								
vorherige Behandlung mit einer anderen DMT																								
Zeit seit der Erstmanifestation																								
Zeit seit der MS-Diagnose																								
Vorjahr																								
Supernode Krankheitsaktivität Baseline																								
Verträglichkeit der Vortherapie																								
Erwartungen an die Behandlung																								
kürzliche Schwangerschaft																								
absolute Lymphozytenzahl																								
CSF Status																								
MS-Zentrum																								
gesundheitsbezogene Lebensqualität																								

kein Zusammenhang, Zusammenhang; Unsicherheit über Zusammenhang;

>: Einfluss von Zeile auf Spalte; <: Einfluss von Spalte auf Zeile; < >: bidirektionaler Pfeil

Supernode Veränderung Krankheitsaktivität im Vorjahr umfasst: EDSS Veränderung im Vorjahr, MSFC Score Veränderung im Vorjahr, Anzahl neuer Läsionen im Vorjahr, neue infratentorielle/spinale Läsionen im Vorjahr, Volumen der Läsionen Veränderung im Vorjahr, Anzahl der vorherigen Schübe, Schwere der (vorherigen) Schübe

Supernode Krankheitsaktivität Baseline umfasst: EDSS Score, MSFC Score, Anzahl der Läsionen, infratentorielle / spinale Läsionen

Abbildung 11: Dokumentation der Zusammenhänge zwischen Therapieentscheidung, Endpunkt und den einzelnen Variablen

E.2.2 DAGitty Modell-Code des finalen DAG (Anordnung im zeitlichen Raster)

```
dag {  
  "Art der beruflichen Tätigkeit" [pos="0.042,-0.066"]  
  "CSF Status" [pos="-1.052,0.631"]  
  "Erwartung an die Behandlung" [latent,pos="0.050,0.690"]  
  "Komorbiditäten" [pos="-0.526,-0.857"]  
  "MS Zentrum" [pos="0.036,-0.867"]  
  "Supernode Krankheitsaktivität Baseline" [pos="0.028,-1.892"]  
  "Supernode Veränderung Krankheitsaktivität im Vorjahr" [pos="-0.522,-1.698"]  
  "Verträglichkeit der Vortherapie" [pos="-0.528,1.168"]  
  "Zeit seit der Erstmanifestation" [pos="-1.178,-0.051"]  
  "Zeit seit der MS Diagnose" [pos="-1.056,0.312"]  
  "absolute Lymphozytenzahl" [pos="0.034,-1.240"]  
  "gesundheitsbezogene Lebensqualität" [pos="0.059,1.059"]  
  "kürzliche Schwangerschaft" [pos="-0.528,0.257"]  
  "vorherige Behandlung mit einer anderen DMT" [pos="-0.528,0.710"]  
  Abstammung [pos="-1.363,-1.210"]  
  Alter [pos="-1.365,-1.579"]  
  BMI [pos="-0.522,-1.235"]  
  Bildungsgrad [pos="-1.297,-0.922"]  
  Endpunkt [outcome,pos="0.562,-0.449"]  
  Fatigue [pos="0.025,-1.579"]  
  Geschlecht [pos="-1.365,-1.902"]  
  Patientenwunsch [latent,pos="0.044,0.287"]  
  Tabakkonsum [pos="-1.292,-0.524"]  
  Therapieentscheidung [exposure,pos="0.036,-0.454"]  
  "Art der beruflichen Tätigkeit" -> "Erwartung an die Behandlung"  
  "Art der beruflichen Tätigkeit" -> "gesundheitsbezogene Lebensqualität"  
  "Art der beruflichen Tätigkeit" -> Patientenwunsch  
  "Art der beruflichen Tätigkeit" -> Therapieentscheidung  
  "CSF Status" -> "Zeit seit der MS Diagnose"  
  "CSF Status" -> Endpunkt  
  "Erwartung an die Behandlung" -> Endpunkt  
  "Erwartung an die Behandlung" -> Therapieentscheidung  
  "Komorbiditäten" -> "Art der beruflichen Tätigkeit"  
  "Komorbiditäten" -> "Erwartung an die Behandlung"  
  "Komorbiditäten" -> "MS Zentrum"  
  "Komorbiditäten" -> "Verträglichkeit der Vortherapie"  
  "Komorbiditäten" -> "absolute Lymphozytenzahl"  
  "Komorbiditäten" -> "gesundheitsbezogene Lebensqualität"  
  "Komorbiditäten" -> BMI  
  "Komorbiditäten" -> Endpunkt  
  "Komorbiditäten" -> Fatigue  
  "Komorbiditäten" -> Patientenwunsch  
  "Komorbiditäten" -> Therapieentscheidung  
  "MS Zentrum" -> "Erwartung an die Behandlung"  
  "MS Zentrum" -> Endpunkt  
  "MS Zentrum" -> Therapieentscheidung  
  "Supernode Krankheitsaktivität Baseline" -> "Erwartung an die Behandlung"  
  "Supernode Krankheitsaktivität Baseline" -> "MS Zentrum"  
  "Supernode Krankheitsaktivität Baseline" -> "gesundheitsbezogene Lebensqualität"  
  "Supernode Krankheitsaktivität Baseline" -> Endpunkt  
  "Supernode Krankheitsaktivität Baseline" -> Patientenwunsch  
  "Supernode Krankheitsaktivität Baseline" -> Therapieentscheidung
```

"Supernode Veränderung Krankheitsaktivität im Vorjahr" -> "Erwartung an die Behandlung"
"Supernode Veränderung Krankheitsaktivität im Vorjahr" -> "MS Zentrum"
"Supernode Veränderung Krankheitsaktivität im Vorjahr" -> "Supernode Krankheitsaktivität Baseline"
"Supernode Veränderung Krankheitsaktivität im Vorjahr" -> "gesundheitsbezogene Lebensqualität"
"Supernode Veränderung Krankheitsaktivität im Vorjahr" -> Endpunkt
"Supernode Veränderung Krankheitsaktivität im Vorjahr" -> Fatigue
"Supernode Veränderung Krankheitsaktivität im Vorjahr" -> Patientenwunsch
"Supernode Veränderung Krankheitsaktivität im Vorjahr" -> Therapieentscheidung
"Verträglichkeit der Vortherapie" -> "Art der beruflichen Tätigkeit"
"Verträglichkeit der Vortherapie" -> "Erwartung an die Behandlung"
"Verträglichkeit der Vortherapie" -> "MS Zentrum"
"Verträglichkeit der Vortherapie" -> "Supernode Krankheitsaktivität Baseline"
"Verträglichkeit der Vortherapie" -> "Supernode Veränderung Krankheitsaktivität im Vorjahr"
"Verträglichkeit der Vortherapie" -> "absolute Lymphozytenzahl"
"Verträglichkeit der Vortherapie" -> "gesundheitsbezogene Lebensqualität"
"Verträglichkeit der Vortherapie" -> Fatigue
"Verträglichkeit der Vortherapie" -> Patientenwunsch
"Verträglichkeit der Vortherapie" -> Therapieentscheidung
"Zeit seit der Erstmanifestation" -> "Art der beruflichen Tätigkeit"
"Zeit seit der Erstmanifestation" -> "Erwartung an die Behandlung"
"Zeit seit der Erstmanifestation" -> "Supernode Krankheitsaktivität Baseline"
"Zeit seit der Erstmanifestation" -> "Supernode Veränderung Krankheitsaktivität im Vorjahr"
"Zeit seit der Erstmanifestation" -> "Verträglichkeit der Vortherapie"
"Zeit seit der Erstmanifestation" -> "Zeit seit der MS Diagnose"
"Zeit seit der Erstmanifestation" -> "gesundheitsbezogene Lebensqualität"
"Zeit seit der Erstmanifestation" -> "vorherige Behandlung mit einer anderen DMT"
"Zeit seit der Erstmanifestation" -> Endpunkt
"Zeit seit der Erstmanifestation" -> Fatigue
"Zeit seit der Erstmanifestation" -> Patientenwunsch
"Zeit seit der Erstmanifestation" -> Therapieentscheidung
"Zeit seit der MS Diagnose" -> "Art der beruflichen Tätigkeit"
"Zeit seit der MS Diagnose" -> "Erwartung an die Behandlung"
"Zeit seit der MS Diagnose" -> "MS Zentrum"
"Zeit seit der MS Diagnose" -> "Supernode Krankheitsaktivität Baseline"
"Zeit seit der MS Diagnose" -> "Supernode Veränderung Krankheitsaktivität im Vorjahr"
"Zeit seit der MS Diagnose" -> "Verträglichkeit der Vortherapie"
"Zeit seit der MS Diagnose" -> "gesundheitsbezogene Lebensqualität"
"Zeit seit der MS Diagnose" -> "kürzliche Schwangerschaft"
"Zeit seit der MS Diagnose" -> "vorherige Behandlung mit einer anderen DMT"
"Zeit seit der MS Diagnose" -> Endpunkt
"Zeit seit der MS Diagnose" -> Fatigue
"Zeit seit der MS Diagnose" -> Patientenwunsch
"Zeit seit der MS Diagnose" -> Therapieentscheidung
"absolute Lymphozytenzahl" -> "Erwartung an die Behandlung"
"absolute Lymphozytenzahl" -> Endpunkt
"absolute Lymphozytenzahl" -> Therapieentscheidung
"absolute Lymphozytenzahl" <-> Fatigue
"gesundheitsbezogene Lebensqualität" -> "Erwartung an die Behandlung"
"gesundheitsbezogene Lebensqualität" -> Endpunkt
"gesundheitsbezogene Lebensqualität" -> Patientenwunsch
"gesundheitsbezogene Lebensqualität" -> Therapieentscheidung
"kürzliche Schwangerschaft" -> "Erwartung an die Behandlung"
"kürzliche Schwangerschaft" -> "Komorbiditäten"
"kürzliche Schwangerschaft" -> "MS Zentrum"
"kürzliche Schwangerschaft" -> "Supernode Krankheitsaktivität Baseline"

"kürzliche Schwangerschaft" -> "Supernode Veränderung Krankheitsaktivität im Vorjahr"
"kürzliche Schwangerschaft" -> "Verträglichkeit der Vortherapie"
"kürzliche Schwangerschaft" -> "absolute Lymphozytenzahl"
"kürzliche Schwangerschaft" -> "gesundheitsbezogene Lebensqualität"
"kürzliche Schwangerschaft" -> "vorherige Behandlung mit einer anderen DMT"
"kürzliche Schwangerschaft" -> BMI
"kürzliche Schwangerschaft" -> Endpunkt
"kürzliche Schwangerschaft" -> Fatigue
"kürzliche Schwangerschaft" -> Patientenwunsch
"kürzliche Schwangerschaft" -> Therapieentscheidung
"vorherige Behandlung mit einer anderen DMT" -> "Erwartung an die Behandlung"
"vorherige Behandlung mit einer anderen DMT" -> "Komorbiditäten"
"vorherige Behandlung mit einer anderen DMT" -> "MS Zentrum"
"vorherige Behandlung mit einer anderen DMT" -> "Supernode Krankheitsaktivität Baseline"
"vorherige Behandlung mit einer anderen DMT" -> "Supernode Veränderung Krankheitsaktivität im Vorjahr"
"vorherige Behandlung mit einer anderen DMT" -> "Verträglichkeit der Vortherapie"
"vorherige Behandlung mit einer anderen DMT" -> "absolute Lymphozytenzahl"
"vorherige Behandlung mit einer anderen DMT" -> "gesundheitsbezogene Lebensqualität"
"vorherige Behandlung mit einer anderen DMT" -> Fatigue
"vorherige Behandlung mit einer anderen DMT" -> Patientenwunsch
"vorherige Behandlung mit einer anderen DMT" -> Therapieentscheidung
Abstammung -> "Art der beruflichen Tätigkeit"
Abstammung -> "Erwartung an die Behandlung"
Abstammung -> "Komorbiditäten"
Abstammung -> "MS Zentrum"
Abstammung -> "Supernode Krankheitsaktivität Baseline"
Abstammung -> "Supernode Veränderung Krankheitsaktivität im Vorjahr"
Abstammung -> "Zeit seit der Erstmanifestation"
Abstammung -> "absolute Lymphozytenzahl"
Abstammung -> "vorherige Behandlung mit einer anderen DMT"
Abstammung -> BMI
Abstammung -> Bildungsgrad
Abstammung -> Endpunkt
Abstammung -> Tabakkonsum
Alter -> "Art der beruflichen Tätigkeit"
Alter -> "Erwartung an die Behandlung"
Alter -> "Komorbiditäten"
Alter -> "Supernode Krankheitsaktivität Baseline"
Alter -> "Supernode Veränderung Krankheitsaktivität im Vorjahr"
Alter -> "Verträglichkeit der Vortherapie"
Alter -> "Zeit seit der Erstmanifestation"
Alter -> "absolute Lymphozytenzahl"
Alter -> "gesundheitsbezogene Lebensqualität"
Alter -> "kürzliche Schwangerschaft"
Alter -> "vorherige Behandlung mit einer anderen DMT"
Alter -> BMI
Alter -> Bildungsgrad
Alter -> Endpunkt
Alter -> Fatigue
Alter -> Patientenwunsch
Alter -> Tabakkonsum
Alter -> Therapieentscheidung
BMI -> "Art der beruflichen Tätigkeit"
BMI -> "Erwartung an die Behandlung"
BMI -> "Supernode Krankheitsaktivität Baseline"

BMI -> "Supernode Veränderung Krankheitsaktivität im Vorjahr"
BMI -> "absolute Lymphozytenzahl"
BMI -> "gesundheitsbezogene Lebensqualität"
BMI -> Endpunkt
BMI -> Fatigue
BMI -> Therapieentscheidung
Bildungsgrad -> "Art der beruflichen Tätigkeit"
Bildungsgrad -> "Erwartung an die Behandlung"
Bildungsgrad -> "Komorbiditäten"
Bildungsgrad -> "MS Zentrum"
Bildungsgrad -> "Zeit seit der MS Diagnose"
Bildungsgrad -> "gesundheitsbezogene Lebensqualität"
Bildungsgrad -> BMI
Bildungsgrad -> Fatigue
Bildungsgrad -> Patientenwunsch
Bildungsgrad -> Tabakkonsum
Bildungsgrad -> Therapieentscheidung
Fatigue -> "Art der beruflichen Tätigkeit"
Fatigue -> "Erwartung an die Behandlung"
Fatigue -> "MS Zentrum"
Fatigue -> "gesundheitsbezogene Lebensqualität"
Fatigue -> Endpunkt
Fatigue -> Patientenwunsch
Geschlecht -> "Art der beruflichen Tätigkeit"
Geschlecht -> "Erwartung an die Behandlung"
Geschlecht -> "Komorbiditäten"
Geschlecht -> "Supernode Krankheitsaktivität Baseline"
Geschlecht -> "Supernode Veränderung Krankheitsaktivität im Vorjahr"
Geschlecht -> "Zeit seit der Erstmanifestation"
Geschlecht -> "Zeit seit der MS Diagnose"
Geschlecht -> "kürzliche Schwangerschaft"
Geschlecht -> BMI
Geschlecht -> Endpunkt
Geschlecht -> Patientenwunsch
Geschlecht -> Therapieentscheidung
Patientenwunsch -> "Erwartung an die Behandlung"
Patientenwunsch -> "MS Zentrum"
Patientenwunsch -> Endpunkt
Patientenwunsch -> Therapieentscheidung
Tabakkonsum -> "Erwartung an die Behandlung"
Tabakkonsum -> "Komorbiditäten"
Tabakkonsum -> "Supernode Krankheitsaktivität Baseline"
Tabakkonsum -> "Supernode Veränderung Krankheitsaktivität im Vorjahr"
Tabakkonsum -> "Zeit seit der Erstmanifestation"
Tabakkonsum -> "Zeit seit der MS Diagnose"
Tabakkonsum -> "absolute Lymphozytenzahl"
Tabakkonsum -> "gesundheitsbezogene Lebensqualität"
Tabakkonsum -> BMI
Tabakkonsum -> Endpunkt
Tabakkonsum -> Fatigue
Therapieentscheidung -> Endpunkt
}

E.3 Tabellarische Darstellung der weggelassenen Pfeile zwischen einzelnen Confoundern

Tabelle 26: Dokumentation weggelassener Pfeile zwischen einzelnen Confoundern
(mehreseitige Tabelle)

Kein Pfeil	zwischen der Variablen (1)	und der Variablen (2)
	Geschlecht	Abstammung
		Alter
		Bildungsgrad
		Tabakkonsum
		Fatigue
		vorherige Behandlung mit einer anderen DMT
		Verträglichkeit der Vortherapie
		absolute Lymphozytenzahl
		MS-Zentrum
		gesundheitsbezogene Lebensqualität
	Abstammung	Alter
		Patientenwunsch
		Fatigue
		Zeit seit der MS-Diagnose
		Verträglichkeit der Vortherapie
		kürzliche Schwangerschaft
		gesundheitsbezogene Lebensqualität
	Alter	MS-Zentrum
		Zeit seit der MS-Diagnose
	Bildungsgrad	vorherige Behandlung mit einer anderen DMT
		Zeit seit der Erstmanifestation
		Supernode Veränderung der Krankheitsaktivität im Vorjahr
		Supernode Krankheitsaktivität Baseline
		Verträglichkeit der Vortherapie
		kürzliche Schwangerschaft
		absolute Lymphozytenzahl
	Art der beruflichen Tätigkeit	Tabakkonsum
		Vorherige Behandlung mit einer anderen DMT
		Supernode Veränderung der Krankheitsaktivität im Vorjahr
		Supernode Krankheitsaktivität Baseline
		kürzliche Schwangerschaft
		absolute Lymphozytenzahl
		MS-Zentrum

Tabelle 26: Dokumentation weggelassener Pfeile zwischen einzelnen Confoundern
(mehrsseitige Tabelle)

Kein Pfeil	zwischen der Variablen (1)	und der Variablen (2)
	Patientenwunsch	Tabakkonsum
		BMI
		absolute Lymphozytenzahl
	Tabakkonsum	vorherige Behandlung mit einer anderen DMT
		Verträglichkeit der Vortherapie
		MS-Zentrum
		kürzliche Schwangerschaft
	BMI	vorherige Behandlung mit einer anderen DMT
		Verträglichkeit der Vortherapie
		Zeit seit der Erstmanifestation
		Zeit seit der MS-Diagnose
		MS-Zentrum
	Fatigue	Supernode Krankheitsaktivität Baseline
	Komorbiditäten	Zeit seit der Erstmanifestation
		Zeit seit der MS-Diagnose
		Supernode Veränderung der Krankheitsaktivität im Vorjahr
		Supernode Krankheitsaktivität Baseline
	Zeit seit der Erstmanifestation	kürzliche Schwangerschaft
		absolute Lymphozytenzahl
		MS-Zentrum
	absolute Lymphozytenzahl	Zeit seit der MS-Diagnose
		Supernode Vorjahr
		Supernode Krankheitsaktivität Baseline
		MS-Zentrum
		gesundheitsbezogene Lebensqualität
	MS-Zentrum	gesundheitsbezogene Lebensqualität

Tabelle 26: Dokumentation weggelassener Pfeile zwischen einzelnen Confoundern
(mehreseitige Tabelle)

Kein Pfeil	zwischen der Variablen (1)	und der Variablen (2)
	CSF Status	Geschlecht
		Abstammung
		Alter
		Bildungsgrad
		Art der beruflichen Tätigkeit
		Patientenwunsch
		Tabakkonsum
		BMI
		Fatigue
		Komorbiditäten
		Zeit seit der Erstmanifestation
		vorherige Behandlung mit einer anderen DMT
		Supernode Veränderung der Krankheitsaktivität im Vorjahr
		Supernode Krankheitsaktivität Baseline
		Verträglichkeit der Vortherapie
		Erwartung an die Behandlung
		kürzliche Schwangerschaft
		absolute Lymphozytenzahl
		MS-Zentrum
		gesundheitsbezogene Lebensqualität
BMI: Body-Mass-Index; DMT: krankheitsmodifizierende Therapie; MS: multiple Sklerose		

E.4 Dokumentation der im Workshop vorgenommenen Änderungen am Confounder-Set aus GA23-02

Tabelle 27: Identifiziertes Set potenzieller Confounder aus GA23-02 inklusive vorgeschlagener Operationalisierung in der Indikation RRMS und den im Workshop vorgenommenen Änderungen durch die klinischen Expertinnen und Experten (mehreseitige Tabelle)

Nr. ^a	potenzieller Confounder ^b	vorgeschlagene Operationalisierung ^b
1	Geschlecht	weiblich vs. männlich vs. divers
2	Alter	stetig
3	EDSS Score	stetig
4	EDSS Veränderung im Vorjahr	stetig
5	MSFC Score	stetig
	MSFC Score Veränderung im Vorjahr	stetig
6	Fatigue	mithilfe eines validierten Instruments: stetig
7	Komorbiditäten	Anzahl der Komorbiditäten: stetig
8	absolute Lymphozytenzahl; Lymphopenie	absolute Lymphozytenzahl: stetig
9	Zentrum für MS	MS-spezialisiertes vs. nicht spezialisiertes Zentrum
10	BMI	stetig
11	Tabakkonsum	ja vs. nein <u>oder</u> Raucher vs. ehemaliger Raucher vs. Nieraucher
12	infratentorielle / spinale Läsionen	ja vs. nein
	neue infratentorielle / spinale Läsionen im Vorjahr	ja vs. nein
13	Anzahl der Läsionen ^c	stetig
	Anzahl neuer Läsionen im Vorjahr^c	stetig
14	Volumen der Läsionen Veränderung im Vorjahr^d	(sichtbare) Vergrößerung der Läsionen im Vorjahr: ja vs. nein
17	Verträglichkeit der Vortherapie	Abbruch wegen UEs: ja vs. nein vs. nicht zutreffend
	Erwartung an die Behandlung	keine adäquate Operationalisierung möglich
18	vorherige Behandlung mit einer anderen DMT	ja vs. nein
19	Patientenwunsch	keine adäquate Operationalisierung möglich
20	gesundheitsbezogene Lebensqualität	mithilfe eines validierten Instruments: stetig
21	Anzahl der vorherigen Schübe	Schübe im letzten Jahr: stetig
22	Schwere der (vorherigen) Schübe	mit Steroiden behandelter Schub im letzten Jahr: ja vs. nein
23	Zeit seit der Diagnose der MS	Zeit seit der MS-Diagnose: stetig

Tabelle 27: Identifiziertes Set potenzieller Confounder aus GA23-02 inklusive vorgeschlagener Operationalisierung in der Indikation RRMS und den im Workshop vorgenommenen Änderungen durch die klinischen Expertinnen und Experten (mehrseitige Tabelle)

Nr. ^a	potenzieller Confounder ^b	vorgeschlagene Operationalisierung ^b
24	Cerebrospinalflüssigkeit (CSF) Status	Liquor spezifische oligoklonale Banden: ja vs. nein
25	Art der beruflichen Tätigkeit	berufliche Tätigkeit, deren Durchführung durch die MS eingeschränkt wird: ja / nein
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
	Zeit seit der Erstmanifestation	Zeit seit Erstmanifestation: stetig
a. Die Nummerierung entspricht der ursprünglich in GA23-02 [8] Tabelle 1 vorgenommenen Nummerierung der 28 zusammengeführten Confounder. Confounder ohne Nummer wurden neu aufgenommen. b. Alle identifizierten potenziellen Confounder-Variablen inklusive der Zuordnung zu den jeweiligen potenziellen Confoundern sind in GA23-02 [8] Anhang B zu entnehmen. Änderungen am ursprünglich (in GA23-02) identifizierten Confounder-Set, deren Operationalisierung oder neue Confounder sind fett hervorgehoben. c. Der Confounder Nr. 13 Anzahl der Kontrastmittel-anreichernden Läsionen und Nr. 15 Anzahl T2-Läsionen wurden zur Anzahl der Läsionen zusammengeführt, da gemäß Leitlinie [75] bei MRT-Verlaufsuntersuchungen auf die Gabe von Gadolinium-haltigem Kontrastmittel verzichtet werden sollte. d. Der Confounder Nr. 14 Volumen T1-hypointenser Läsionen und Nr. 16 Volumen der T2-Läsionen wurden zum Volumen der Läsionen Veränderung im Vorjahr zusammengeführt. BMI: Body-Mass-Index; DMT: krankheitsmodifizierende Therapie; EDSS: Expanded Disability Status Scale; MS: multiple Sklerose; MSFC: Multiple Sclerosis Functional Composite; UEs: unerwünschte Ereignisse		

E.5 Dokumentation der im Nachgang zum Workshop vorgenommenen Anpassungen der Zusammenhänge

Tabelle 28: Änderung der besprochenen Zusammenhänge zwischen den einzelnen Confoundern im Nachgang des Workshops^a

Einfluss von der Variablen	Im Workshop besprochener Zusammenhang	Änderung des Zusammenhangs	Einfluss zur Variablen
BMI	<	>	Art der beruflichen Tätigkeit
Komorbidität	<	>	BMI
Zeit seit der MS-Diagnose	<	>	MS-Zentrum
Zeit seit der MS-Diagnose	<	>	kürzliche Schwangerschaft (in den letzten 3 Monaten)
Supernode Krankheitsaktivität Baseline	>		Fatigue
BMI	>		Zeit seit der MS-Diagnose
BMI	>		Zeit seit der Erstmanifestation
kürzliche Schwangerschaft	>		Tabakkonsum
>: Einfluss von Spalte 1 auf Spalte 2 <: Einfluss von Spalte 2 auf Spalte 1 a. Aufgrund von Zyklen im DAG war im Nachgang zur Workshop-Fortsetzung eine erneute Überprüfung der Zusammenhänge zwischen einzelnen Confoundern notwendig. Unter Berücksichtigung des zeitlichen Aspekts, wann die Variablen erhoben werden, wurden einzelne Zusammenhänge geändert. Ein grünes Rechteck zeigt einen Zusammenhang an, wobei der Pfeil die Richtung des Zusammenhangs vorgibt. Ein rotes Rechteck deutet keinen Zusammenhang an. BMI: Body-Mass-Index; DAG: gerichteter azyklischer Graph; MS: multiple Sklerose			

Anhang F Online-Umfrage zur Bewertung der Relevanz der Confounder

F.1 IQWiG-Begleitschreiben zur Online-Umfrage

Umfrage zum Einfluss von Confoundern auf die Therapieentscheidung und Endpunkte in der Indikation RRMS

Das IQWiG erarbeitet in mehreren wissenschaftlichen Projekten eine Methodik zur systematischen Identifikation von Confoundern für nicht randomisierte vergleichende Studien (Beobachtungsstudien). Confounder (Störgrößen) sind Faktoren, die sowohl die Therapieentscheidung als auch den Studien-Endpunkt beeinflussen und somit in einer Beobachtungsstudie den untersuchten Behandlungseffekt verfälschen können. Die Berücksichtigung solcher Faktoren (Confounder) in Beobachtungsstudien ist von großer Bedeutung, um möglichst aussagekräftige Ergebnisse zu erhalten.

Die Ergebnisse des ersten Projekts sind im Arbeitspapier GA23-02 „Systematische Confounderidentifikation in der Indikation schubförmig remittierende multiple Sklerose (RRMS)“ veröffentlicht ([hier](#) verfügbar). In GA23-02 wurden durch eine externe Arbeitsgruppe über eine systematische Literaturrecherche und Experteninterviews mögliche Confounder in der Indikation RRMS identifiziert. Das Ziel war es, die identifizierten Confounder für die Nachbildung der randomisierten kontrollierten Studie (RCT) CONFIRM ([hier](#)) zum Vergleich der Interventionen Dimethylfumarat vs. Glatirameracetat anhand eines nicht randomisierten Vergleichs zu nutzen¹. Als Endpunkte wurde die Schubrate, die Behinderungsprogression, die gesundheitsbezogene Lebensqualität und Nebenwirkungen ausgewählt. Das Ergebnis des Projekts GA23-02 war zum einen ein Vorschlag für ein 3-schrittiges Vorgehen zur Identifikation von Confoundern und zum anderen, bezogen auf das Beispiel RRMS, eine Liste von 28 möglichen Confoundern.

In einem aktuellen Projekt (GA25-02) evaluiert das IQWiG Maßnahmen zur Aufwandsreduktion bei der systematischen Identifikation und der Erhebung von Confoundern ([hier](#)). Aufbauend auf dem Ergebnis des vorherigen Projekts wurde bereits ein Workshop mit Klinikerinnen und Klinikern zur Validierung der relevanten Confounder für die vorliegende Fragestellung in der Indikation RRMS durchgeführt.

Zur Identifikation möglicher Ansatzpunkte, die Liste der relevanten Confounder (n = 31) noch weiter zu reduzieren, möchten wir im jetzigen Schritt die Stärke des Einflusses der jeweiligen

¹ Die RCT CONFIRM wurde bereits vor einigen Jahren hierfür ausgewählt, da die RCT abgeschlossen ist, die beiden Therapien seit mehreren Jahren im untersuchten Anwendungsgebiet zugelassen und zudem in Deutschland verfügbar sind. Darüber hinaus sind sowohl das Studienprotokoll als auch die Ergebnisse öffentlich verfügbar. Auch wenn die Fragestellung nicht mehr den aktuellen Versorgungsstandard abbildet, ist sie ein geeigneter Ausgangspunkt für ein Modellprojekt zur Erarbeitung des methodischen Vorgehens für eine adäquate Confounderidentifikation und -erhebung.

Confounder auf Therapieentscheidung und / oder Endpunkt bewerten. Hierzu sind wir auf Ihre Unterstützung angewiesen.

F.2 Online-Umfrage

Confounder in der Indikation RRMS - Befragung von Expertinnen und Experten

Willkommen zur Umfrage des IQWiG

Derzeit evaluiert das IQWiG im Rahmen einer wissenschaftlichen Ausarbeitung Maßnahmen zur Aufwandsreduktion bei der systematischen Identifikation und der Erhebung von Confoundern ([hier](#)). Vielen Dank, dass Sie uns bei diesem Projekt unterstützen möchten.

Nähere Informationen entnehmen Sie bitte dem Anschreiben.

Wir bitten Sie nun um Ihre fachliche Einschätzung, um die Stärke des Einflusses der in der Indikation RRMS identifizierten Confounder zu bewerten. Bitte geben Sie auf einer 5-Punkte-Skala von „1 – sehr geringer Einfluss“ bis „5 – sehr starker Einfluss“ an, inwieweit jeder Confounder (I) Ihre Therapieentscheidung und (II) die Endpunkte (Schubrate, Behinderungsprogression, gesundheitsbezogene Lebensqualität oder unerwünschte Ereignisse) beeinflusst.

Das Ausfüllen dieser Umfrage dauert etwa 20 Minuten.

Vielen Dank für Ihre Teilnahme. Wir wissen Ihre Unterstützung sehr zu schätzen.

Informationen zum Datenschutz

Es erfolgt keine Erfassung personenbezogener Daten. Ihre Antworten werden vollständig anonym erhoben und ausgewertet. Es werden auch keine Systemdaten gespeichert, also weder Ihre IP-Adresse noch technische Metadaten erfasst oder ausgewertet. Die Umfrage ist zudem so gestaltet, dass auch keine Rückschlüsse auf Ihre Person durch die Antworten möglich sind. Die Umfrage wird über ein DSGVO-konformes Tool durchgeführt, dessen Server in der EU stehen. Ihre Antworten werden verschlüsselt übertragen.

Die Umfrage wird nach Export der Ergebnisse gelöscht. Ihre Antworten werden für die Datenanalyse verwendet und die generierten Ergebnisse veröffentlicht.

Dies ist eine anonyme Umfrage.

In den Umfrageantworten werden keine persönlichen Informationen über Sie gespeichert, es sei denn, in einer Frage wird explizit danach gefragt.

Wenn Sie für diese Umfrage einen Zugangscode benutzt haben, so können Sie sicher sein, dass der Zugangsschlüssel nicht zusammen mit den Daten abgespeichert wurde. Er wird in einer getrennten Tabelle aufbewahrt und nur aktualisiert, um zu speichern, ob Sie diese Umfrage abgeschlossen haben oder nicht. Es gibt keinen Weg, die Zugangscode mit den Umfrageergebnissen zusammenzuführen.

Übersicht über die Confounder inkl. Operationalisierung

Um Ihnen die Beurteilung zu erleichtern, haben wir die Confounder bereits im Vorfeld in verschiedene thematische Gruppen unterteilt. Bitte schauen Sie sich zuerst die gesamte Liste der 31 Confounder inklusive der entsprechenden Operationalisierung an.

Sofern nicht anders angegeben, soll die Ausprägung des jeweiligen Confounders zum Zeitpunkt der Therapieentscheidung berücksichtigt werden.

Faktoren	Operationalisierung
Demographisch	
Geschlecht	weiblich vs. männlich vs. divers
Alter	in Jahren
Abstammung	weiß, asiatisch, schwarz oder afroamerikanisch, andere
Sozioökonomisch	
Bildungsgrad	Bildungszeit in Jahren
Art der beruflichen Tätigkeit	berufliche Tätigkeit, deren Durchführung durch die MS eingeschränkt wird: ja/nein
Lebensstil und Komorbiditäten	
Tabakkonsum	Raucher vs. ehemaliger Raucher vs. Nieraucher
Body-Mass-Index	kg/m ²
Komorbiditäten	Anzahl der Komorbiditäten (wie bspw. Depression, Allergie, kardiovaskuläre Komorbiditäten, chronische Infektionen, dermatologische Erkrankungen, Diabetes, Kopfschmerzen, Hautprobleme, Fehlfunktion der Schilddrüse, gastrointestinale Erkrankungen)
Krankheitsdauer und Vortherapie	
Zeit seit der Erstmanifestation	in Jahren
Zeit seit der MS-Diagnose	in Jahren
Vorherige Behandlung mit einer anderen krankheitsmodifizierenden Therapie	ja vs. nein
Verträglichkeit der Vortherapie	Abbruch wegen unerwünschten Ereignissen: ja vs. nein vs. nicht zutreffend

Krankheitsschwere – klinisch	
EDSS Veränderung im Vorjahr	Veränderung im Vorjahr
EDSS Score	absolut
Fatigue	Punktwert, abhängig vom verwendeten Messinstrument, z. B. auf einer Skala von 0 bis 100
MSFC Score Veränderung im Vorjahr	Veränderung im Vorjahr
MSFC Score	absolut
Anzahl der vorherigen Schübe	Schübe im letzten Jahr
Schwere der vorherigen Schübe	mit Steroiden behandelter Schub im letzten Jahr: ja vs. nein
Krankheitsschwere – radiologisch	
Anzahl neuer Läsionen im Vorjahr	Anzahl neuer Läsionen im Vergleich zum Vorjahr
Anzahl der Läsionen	Anzahl
Neue Infratentorielle/spinale Läsionen im Vorjahr	Neue infratentorielle/spinale Läsionen im Vergleich zum Vorjahr: ja vs. nein
Infratentorielle/spinale Läsionen	ja vs. nein
Volumen der Läsionen Veränderung im Vorjahr	(sichtbare) Vergrößerung der Läsionen im Vorjahr: ja vs. nein
Sonstiges	
MS-Zentrum	MS-spezialisiertes vs. nicht spezialisiertes Zentrum
CSF Status	Liquor spezifische oligoklonale Banden: ja vs. nein
Absolute Lymphozytenzahl	Zellen/ μ l
Gesundheitsbezogene Lebensqualität	Punktwert, abhängig vom verwendeten Messinstrument, z. B. auf einer Skala von 0 bis 100
Kürzliche Schwangerschaft (in den letzten 3 Monaten)	ja vs. nein
Patientenwunsch	z. B. Präferenzen hinsichtlich Verabreichungsform, Bedenken bezüglich Nebenwirkungen
Erwartung an die Behandlung	Perspektive der behandelnden Person

Bewerten Sie nun **im Kontext aller identifizierten Confounder** den entsprechenden Einfluss des jeweiligen Confounders auf die Therapieentscheidung und auf die unten genannten Endpunkte. Achten Sie dabei auch auf die jeweilige Operationalisierung der Confounder.

Versetzen Sie sich hierzu in die Situation, dass Sie für eine*n Ihrer erwachsenen Patient*innen mit RRMS die Entscheidung treffen müssen, ob Sie Dimethylfumarat oder Glatirameracetat verordnen. Gehen Sie hierfür davon aus, dass Ihnen zu jeder Patientin / jedem Patienten zu allen Confoundern Informationen vorliegen. Da das Ziel dieser Umfrage ist, die einzelnen Confounder hinsichtlich ihrer Relevanz für die Therapieentscheidung bzw. die verschiedenen Endpunkte zu differenzieren, ist es hilfreich für uns, wenn Sie wann immer möglich die **gesamte Spanne der 5 Kategorien** von „1 – sehr geringer Einfluss“ bis „5 – sehr starker Einfluss“ ausschöpfen. Vielen Dank.

	Wie stark wird Ihre Therapieentscheidung, ob Sie Dimethylfumarat oder Glatirameracetat verordnen, von den folgenden Confoundern beeinflusst?					Wie stark werden die Endpunkte (Schubrate, Behinderungsprogression, gesundheitsbezogene Lebensqualität oder unerwünschte Ereignisse) durch folgende Confounder beeinflusst? Ggf. unterscheidet sich Ihre Einschätzung je Endpunkt. Geben Sie dann jeweils die Einschätzung für den Endpunkt, der am stärksten beeinflusst wird, an.				
	sehr geringer Einfluss 1	2	3	4	sehr starker Einfluss 5	sehr geringer Einfluss 1	2	3	4	sehr starker Einfluss 5
Demographie - Geschlecht (männlich / weiblich / divers)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Alter (in Jahren)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Abstammung (weiß, asiatisch, schwarz oder afroamerikanisch, andere)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Sozioökonomisch - Bildungsgrad (Bildungszeit in Jahren)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Art der beruflichen Tätigkeit (Tätigkeit durch MS eingeschränkt: ja/ nein)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Lebensstil und Komorbiditäten - Tabakkonsum (Raucher*in: aktiv/ehemals/nie)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Body-Mass-Index (kg/m²)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Komorbiditäten (Anzahl der Komorbiditäten)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Krankheitsdauer und Vortherapie - Zeit seit der Erstmanifestation (in Jahren)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Zeit seit der MS-Diagnose (in Jahren)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Vorherige Behandlung mit einer anderen krankheitsmodifizierenden Therapie (ja / nein)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Verträglichkeit der Vortherapie (Abbruch wegen UE: ja / nein / nicht zutreffend)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

	sehr geringer Einfluss 1	2	3	4	sehr starker Einfluss 5		sehr geringer Einfluss 1	2	3	4	sehr starker Einfluss 5
Krankheitsschwere – klinisch - EDSS Veränderung im Vorjahr (Veränderung im Vorjahr)	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐
EDSS Score (absolut)	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐
Fatigue (Abhängig vom verwendeten Messinstrument)	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐
MSFC Score Veränderung im Vorjahr (Veränderung im Vorjahr)	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐
MSFC Score (absolut)	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐
Anzahl der vorherigen Schübe (Schübe im letzten Jahr)	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐
Schwere der vorherigen Schübe (mit Steroiden behandelter Schub im letzten Jahr: ja/mein)	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐

Krankheitsschwere – radiologisch - Anzahl neuer Läsionen im Vorjahr (Anzahl)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Anzahl der Läsionen (Anzahl)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Neue infratentorielle/spinale Läsio- nen im Vorjahr (Neue infratentorielle/spinale Läsionen im Vergleich zum Vor- jahr: ja/mein)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Infratentorielle/spinale Läsionen (ja/mein)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

	sehr geringer Einfluss 1	2	3	4	sehr starker Einfluss 5		sehr geringer Einfluss 1	2	3	4	sehr starker Einfluss 5
Volumen der Läsionen Veränderung im Vorjahr ((sichtbare) Vergrößerung der Läsionen im Vorjahr: ja/nein)	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐
Sonstiges - MS-Zentrum (MS-spezialisiertes vs. nicht spezialisiertes Zentrum)	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐
CSF Status (Liquor spezifische oligoklonale Banden: ja/nein)	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐
Absolute Lymphozytenzahl (Zellen/μl)	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐
Gesundheitsbezogene Lebensquali- tät (Abhängig vom verwendeten Messinstrument)	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐
Kürzliche Schwangerschaft (in den letzten 3 Monaten) (ja/nein)	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐
Patientenwunsch (z. B. Präferenzen hinsichtlich Verabreichungsform, Bedenken bezüglich Nebenwirkungen)	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐
Erwartung an die Behandlung (Perspektive der behandelnden Person)	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐

F.3 Ergebnis der Online-Umfrage

Tabelle 29: Ergebnis der Online-Umfrage (n = 18) zur Bewertung der Stärke des Einflusses aller einzelnen Confounder auf die Therapieentscheidung bzw. auf die Endpunkte (mehreseitige Tabelle)

Confounder	Stärke des Einflusses des Confounders auf:						
	Therapieentscheidung		Endpunkte ^a		Kombiniert – Gewichtung 1:1	Kombiniert – Gewichtung 2:3	Kombiniert – Gewichtung 1:2
	MW	Median [Q1; Q3]	MW	Median [Q1; Q3]	Median [Q1; Q3]	Median [Q1; Q3]	Median [Q1; Q3]
neue infratentorielle / spinale Läsionen im Vorjahr	3,89	4,5 [3,0; 5,0]	4,56	5,0 [5,0; 5,0]	4,5 [3,6; 5,0]	4,5 [3,9; 5,0]	4,5 [4,0; 5,0]
EDSS Veränderung im Vorjahr	3,89	4,0 [3,0; 5,0]	4,50	5,0 [5,0; 5,0]	4,5 [3,6; 5,0]	4,6 [3,9; 5,0]	4,7 [4,0; 5,0]
infratentorielle / spinale Läsionen	3,89	4,5 [3,0; 5,0]	4,39	5,0 [4,0; 5,0]	4,5 [3,5; 5,0]	4,4 [3,7; 5,0]	4,3 [3,8; 5,0]
Schwere der vorherigen Schübe	3,72	4,5 [3,0; 5,0]	4,28	5,0 [4,0; 5,0]	4,5 [3,0; 5,0]	4,6 [3,1; 5,0]	4,7 [3,2; 5,0]
Anzahl der vorherigen Schübe	3,67	4,5 [2,3; 5,0]	4,17	5,0 [4,0; 5,0]	4,5 [2,6; 5,0]	4,6 [2,9; 5,0]	4,7 [3,0; 5,0]
Vorherige Behandlung mit einer anderen DMT	4,06	4,5 [3,0; 5,0]	4,17	4,5 [4,0; 5,0]	4,3 [3,5; 5,0]	4,3 [3,6; 5,0]	4,3 [3,7; 5,0]
Anzahl neuer Läsionen im Vorjahr	3,83	4,5 [3,0; 5,0]	4,33	5,0 [4,0; 5,0]	4,3 [3,3; 5,0]	4,3 [3,6; 5,0]	4,3 [3,8; 5,0]
Komorbiditäten	3,78	4 [3,3; 5,0]	3,78	4,0 [3,0; 5,0]	4,0 [3,1; 4,5]	4,0 [3,4; 4,6]	4,0 [3,3; 4,6]
Volumen der Läsionen Veränderung im Vorjahr	3,44	3,5 [3,0; 4,8]	4,06	4,0 [4,0; 5,0]	4,0 [3,0; 4,4]	4,0 [3,0; 4,4]	4,0 [3,0; 4,3]
Verträglichkeit der Vortherapie	3,67	4,5 [3,0; 5,0]	3,39	4,0 [2,3; 4,8]	4,0 [2,6; 4,4]	3,9 [2,9; 4,4]	3,8 [3,0; 4,3]
Patientenwunsch	4,61	5,0 [4,0; 5,0]	2,89	3,0 [2,0; 4,0]	3,8 [3,5; 4,0]	3,6 [3,2; 4,0]	3,5 [3,0; 4,0]
Anzahl der Läsionen	3,39	4,0 [2,3; 4,0]	3,78	4,0 [3,3; 4,0]	3,8 [3,1; 4,0]	3,8 [3,3; 4,2]	3,8 [3,3; 4,3]
Zeit seit der Erstmanifestation	3,28	3,0 [2,3; 4,0]	3,78	4,0 [3,0; 4,0]	3,5 [3,0; 4,0]	3,6 [3,0; 4,0]	3,7 [3,0; 4,0]
EDSS Score	3,17	3,0 [2,3; 4,0]	3,78	4,0 [3,0; 5,0]	3,5 [2,6; 4,5]	3,7 [2,9; 4,6]	3,8 [3,0; 4,7]
Erwartung an die Behandlung	3,94	4,0 [4,0; 5,0]	2,78	3,0 [2,0; 4,0]	3,5 [3,1; 4,0]	3,4 [2,9; 4,0]	3,3 [2,8; 3,9]
Zeit seit der MS-Diagnose	3,11	3,0 [2,3; 4,0]	3,59	4,0 [3,0; 4,0]	3,5 [3,0; 4,0]	3,4 [3,0; 4,0]	3,3 [3,0; 4,0]
MSFC Score Veränderung im Vorjahr	3,00	3,0 [3,0; 3,8]	3,53	4,0 [3,0; 4,0]	3,5 [2,5; 4,0]	3,6 [2,6; 4,0]	3,7 [2,7; 4,0]

Tabelle 29: Ergebnis der Online-Umfrage (n = 18) zur Bewertung der Stärke des Einflusses aller einzelnen Confounder auf die Therapieentscheidung bzw. auf die Endpunkte (mehrsseitige Tabelle)

Confounder	Stärke des Einflusses des Confounders auf:						
	Therapieentscheidung		Endpunkte ^a		Kombiniert – Gewichtung 1:1	Kombiniert – Gewichtung 2:3	Kombiniert – Gewichtung 1:2
	MW	Median [Q1; Q3]	MW	Median [Q1; Q3]	Median [Q1; Q3]	Median [Q1; Q3]	Median [Q1; Q3]
kürzliche Schwangerschaft (in den letzten 3 Monaten)	3,28	3,0 [3,0; 4,0]	3,00	3,0 [2,3; 4,0]	3,5 [2,3; 4,0]	3,3 [2,4; 4,0]	3,2 [2,5; 3,9]
Alter	2,72	3,0 [2,0; 3,8]	3,39	4,0 [3,0; 4,0]	3,5 [2,1; 4,0]	3,6 [2,2; 4,0]	3,7 [2,3; 3,9]
gesundheitsbezogene Lebensqualität	3,12	3,0 [2,0; 4,0]	2,94	3,0 [2,0; 4,0]	3,0 [2,0; 4,0]	3,0 [2,0; 4,0]	3,0 [2,0; 4,0]
MSFC Score	2,78	3,0 [2,0; 3,0]	3,28	3,0 [3,0; 4,0]	3,0 [2,1; 3,5]	3,0 [2,3; 3,6]	3,0 [2,4; 3,7]
MS-Zentrum	2,72	3,0 [1,3; 3,8]	3,12	3,0 [3,0; 4,0]	3,0 [2,5; 3,5]	3,0 [2,8; 3,8]	3,0 [3,0; 4,0]
CSF Status	2,61	3,0 [1,3; 3,8]	3,18	3,0 [3,0; 4,0]	3,0 [2,5; 3,5]	3,0 [2,4; 3,6]	3,0 [2,3; 3,7]
absolute Lymphozytenzahl	2,94	4,0 [1,0; 5,0]	2,35	2,0 [1,0; 3,0]	3,0 [1,5; 3,0]	2,6 [1,4; 3,0]	2,3 [1,3; 3,0]
Fatigue	2,50	2,0 [2,0; 3,0]	3,11	3,0 [2,0; 4,8]	2,8 [2,0; 3,5]	2,9 [2,0; 3,8]	3,0 [2,0; 3,9]
Body-Mass-Index	2,22	2,0 [1,0; 3,0]	3,06	3,0 [2,0; 4,0]	2,8 [2,0; 3,0]	2,9 [2,1; 3,4]	3,0 [2,1; 3,6]
Tabakkonsum	1,83	1,0 [1,0; 2,8]	3,44	4,0 [3,0; 4,0]	2,5 [2,0; 3,0]	2,8 [2,2; 3,3]	3,0 [2,3; 3,5]
Geschlecht	2,33	2,5 [1,0; 3,0]	2,33	2,0 [1,0; 3,0]	2,5 [1,5; 3,0]	2,3 [1,5; 3,0]	2,3 [1,4; 3,0]
Art der beruflichen Tätigkeit	2,11	2,0 [1,0; 3,0]	2,06	1,5 [1,0; 3,0]	2,0 [1,0; 2,9]	1,9 [1,0; 2,9]	1,8 [1,0; 2,8]
Bildungsgrad	1,72	1,0 [1,0; 2,0]	1,94	2,0 [1,0; 3,0]	1,5 [1,0; 2,0]	1,6 [1,0; 2,2]	1,7 [1,0; 2,3]
Abstammung	1,50	1,0 [1,0; 1,0]	1,83	1,0 [1,0; 2,0]	1,0 [1,0; 1,9]	1,0 [1,0; 1,9]	1,0 [1,0; 1,9]
a. Betrachtet wurden die Endpunkte Schubrate, Behinderungsprogression, gesundheitsbezogene Lebensqualität und unerwünschte Ereignisse. Falls die Einschätzung für die einzelnen Endpunkte unterschiedlich ausfiel, sollte die Einschätzung für den Endpunkt angegeben werden, der am stärksten beeinflusst wird. CSF: Cerebrospinalflüssigkeit; DMT: krankheitsmodifizierende Therapie; EDSS: Expanded Disability Status Scale; MS: multiple Sklerose; MSFC: Multiple Sclerosis Functional Composite; MW: Mittelwert; Q1: 1. Quartil; Q3: 3. Quartil							

Anhang G Vorschlag für ein Formblatt zur Offenlegung von Beziehungen bei der Einbindung von Expertinnen und Experten in die Confounderidentifikation

Offenlegung von Beziehungen bei der Einbindung in die Confounderidentifikation

Hintergrund

Eine Confounderidentifikation sollte fachlich unabhängig und vorab im Rahmen der Studienplanung erfolgen. Zu diesem Zweck werden alle in die Confounderidentifikation eingebundenen Personen gebeten, Beziehungen, die das Ziel einer unvoreingenommenen und objektiven Mitwirkung beeinflussen oder potenziell beeinflussen können, sogenannte Interessenkonflikte, offenzulegen. Erfolgt die Confounderidentifikation konkret für die Bewertung eines Arzneimittels (wie z. B. im Rahmen der anwendungsbegleitenden Datenerhebung), geht es insbesondere um Beziehungen zu dem Hersteller des betreffenden Arzneimittels, aber auch um sonstige Beziehungen, die Interessenkonflikte begründen können.

Die Offenlegung solcher Beziehungen erfolgt anhand der folgenden Fragen.

Datenschutz

Die Einzelheiten Ihrer Offenlegung werden vertraulich behandelt. Einblick bekommen nur an der Sichtung beteiligte Mitarbeiterinnen und Mitarbeiter.

Projekt

Betreffendes Arzneimittel² (sofern zutreffend):

Hersteller (sofern zutreffend):

Indikation:

Kontaktdaten

Ggf. Titel, Vorname, Nachname:

² Wenn die Confounderidentifikation konkret für die Bewertung eines Arzneimittels (wie z. B. im Rahmen der anwendungsbegleitenden Datenerhebung) erfolgt, ist dies als betreffendes Arzneimittel anzugeben. Ist dies nicht der Fall, sind die Fragen 1 bis 3 als „nicht zutreffend“ anzukreuzen.

Anstellung / Beratung	Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor bei dem Hersteller des betreffenden Arzneimittels angestellt oder für diesen selbstständig oder ehrenamtlich tätig oder haben Sie diesen 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)?	☐ nein ☐ ja ☐ nicht zutreffend ^a
Honorare / Drittmittel / Sonstige Unterstützung	Haben Sie oder hat Ihr Arbeitgeber bzw. Ihre Praxis oder die Institution, für die Sie tätig sind, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor direkt oder indirekt von dem Hersteller des betreffenden Arzneimittels Honorare (z. B. für Vorträge, Schulungstätigkeiten, Stellungnahmen oder Artikel), sogenannte Drittmittel (d. h. finanzielle Unterstützung z. B. für Forschungsaktivitäten, die Durchführung klinischer Studien, andere wissenschaftliche Leistungen oder Patentanmeldungen) oder sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Unterstützung bei der Ausrichtung einer Veranstaltung, Übernahme von Reisekosten oder Teilnahmegebühren für Fortbildungen / Kongresse) erhalten?	☐ nein ☐ ja ☐ nicht zutreffend ^a
Aktien	Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile des Herstellers des betreffenden Arzneimittels? Besitzen Sie Anteile eines sogenannten Branchenfonds, der auf den Hersteller des betreffenden Arzneimittels ausgerichtet ist?	☐ nein ☐ ja ☐ nicht zutreffend ^a
Sonstiges	Haben Sie andere Interessen, die die Unbefangenheit in Bezug auf die Confounderidentifikation infrage stellen können?	☐ nein ☐ ja
a. nicht zutreffend, wenn oben kein betreffendes Arzneimittel angegeben ist.		

Datum

Unterschrift

Anhang H Handlungsempfehlungen – Checkliste

Handlungsempfehlungen, die aus den Arbeitspapieren GA23-02 [8] und GA25-02 abgeleitet wurden.

Schritt 1: Identifikation

- ☐ Im Sinne der Aufwandsreduktion können die in hochwertigen systematischen Übersichten enthaltenen Primärpublikationen für die Extraktion potenzieller Confounder-Variablen verwendet werden, sofern durch die systematischen Übersichten die gesamte Fragestellung (inkl. Studientyp) abgedeckt wird.
- ☐ Bei der systematischen Informationsbeschaffung sollte zuerst auf Primärpublikationen zu RCTs und Kohortenstudien zurückgegriffen werden. Nachfolgend können auch Leitlinien oder 1-armige Studien berücksichtigt werden.
- ☐ Der Recherchezeitpunkt nach systematischen Übersichten bzw. Primärstudien sollte nicht älter als 12 Monate sein.
- ☐ 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 Sättigungskriteriums von 10 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 (übernommen aus GA23-02 [8])

- ☐ 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: Reduktion

Teilschritt 3a: Identifikation eines minimal ausreichenden Adjustierungssets

- ☐ Für einen systematischen und transparenten Ansatz der Reduktion zu erhebender Confounder sollten unter Einbindung klinischer Expertinnen und Experten ein oder mehrere minimal ausreichende Adjustierungssets mittels gerichteter azyklischer Graphen erarbeitet werden. Es wird empfohlen, diese in einem Präsenz-Workshop mit einem Team aus mindestens 1 methodischen und 3 klinischen Expertinnen und Experten zu erarbeiten.
- ☐ Weggelassene Pfeile zur Therapieentscheidung und zu Endpunkten müssen begründet werden. Annahmen für einen fehlenden kausalen Zusammenhang zwischen den einzelnen Confoundern müssen mindestens in tabellarischer Form transparent dokumentiert werden. Im Bericht zur Confounderidentifikation muss das finale DAG bzw. die finalen DAG inkl. Code der verwendeten Software und Matrix abgebildet werden.
- ☐ Die Workshop-Teilnehmenden sollten namentlich benannt und potenzielle Interessenkonflikte offengelegt werden.

Teilschritt 3b: Relevanz der Confounder – Bewertung der Stärke des Einflusses auf die Effektschätzung (optional)

- ☐ Für die weitere Reduktion des finalen Confounder-Sets kann die Stärke des Einflusses sowohl auf die Therapieentscheidung als auch auf die Endpunkte einbezogen werden. Diese kann mithilfe einer Umfrage bei Klinikern und Klinikern ermittelt werden.

Erforderliche Dokumentationen

Literaturrecherche

- ☐ Rationale des Vorgehens bei der Informationsbeschaffung (z. B. systematische Übersichten oder Primärpublikationen, Selektionskriterien) inkl. Angabe des Suchdatums
- ☐ Angabe der Suchstrategie und der durchsuchten Quellen
- ☐ Eindeutige Benennung der Ein- und Ausschlusskriterien
- ☐ Angabe eines Flussdiagramms für den Selektionsprozess
- ☐ Darstellung der Bewertung der Qualität der berücksichtigten systematischen Übersichten (falls zutreffend) (z. B. über AMSTAR-2 [30] oder ROBIS [100])
- ☐ Angabe der eingeschlossenen Referenzen (sowohl bei den systematischen Übersichten als auch Primärpublikationen)

- ☐ 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) 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

Identifikation eines minimal ausreichenden Adjustierungssets

- ☐ Begründung weggelassener Pfeile zur Therapieentscheidung und zu Endpunkten
- ☐ tabellarische Dokumentation der weggelassenen Pfeile zwischen den einzelnen Confoundern
- ☐ Darstellung des finalen DAG bzw. der finalen DAGs inkl. DAGitty Model Code und Matrix zur Dokumentation der Zusammenhänge
- ☐ Benennung von Namen, Fachgebiet und Institution der eingebundenen Expertinnen und Experten und Offenlegung potenzieller Interessenkonflikte (siehe Anhang G)

Umfrage zum Ausschluss von Confoundern untergeordneter Relevanz (optional)

- ☐ Benennung der Anzahl eingebundener klinischer Expertinnen und Experten und Beschreibung wie die Akquise dieser erfolgte
- ☐ Begründung, warum eine Umfrage durchgeführt wurde und ein zusätzlicher Ausschluss von Confoundern über die Relevanzbewertung notwendig war
- ☐ Benennung des Cut-offs und ob dieser a priori definiert war
- ☐ Benennung welche Confounder aus dem finalen Confounder-Set entfernt wurden

Veröffentlichung des Berichts

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

Anhang I 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
Didelez, Vanessa	nein	nein	nein	nein	ja	nein	nein
Rohmann, Jessica	ja	ja	nein	ja	nein	nein	nein

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?