Impact of disease-modifying therapies on MRI and neurocognitive outcomes in relapsing–remitting multiple sclerosis: a protocol for a systematic review and network meta-analysis

Samuel Lees,1 Mathew Dicker,1 Jie En Ku,1 Varun Chaganti,1 Matthew Mew-Sum,1 Nick Wang,1 Angela Smith,2 Christopher Oldmeadow,3 Wooi Lynn Goon,1 Marc Bevan,1 Danielle Lang1,3 Madeleine Hinwood1,3

ABSTRACT

Introduction Disease-modifying therapies (DMTs) are the mainstay of treatment for relapsing–remitting multiple sclerosis (RRMS). There is established evidence that DMTs are effective at reducing relapse rate and disease progression in RRMS, but there has been less consideration to the synthesis of MRI and neurocognitive outcomes, which play an increasingly important role in treatment decisions. The aim of this systematic review and network meta-analysis is to examine the relative efficacy, acceptability and tolerability of DMTs for RRMS, using MRI and neurocognitive outcomes.

Methods and analysis We will search electronic databases, including MEDLINE, Embase and the Cochrane Central Register of Controlled Trials, with no date restrictions. We will also search the websites of international regulatory bodies for pharmaceuticals and international trial registries. We will include parallel group randomised controlled trials of DMTs including interferon beta-1a intramuscular, interferon beta-1a subcutaneous, interferon beta-1b, peginterferon beta-1a, glatiramer acetate, natalizumab, ocrelizumab, alemtuzumab, dimethyl fumarate, teriflunomide, fingolimod, cladribine, ozanimod, mitoxantrone and rituximab, either head-to-head or against placebo in adults with RRMS. Primary outcomes include efficacy (MRI outcomes including new areas on MRI and neurocognitive outcomes) and acceptability (all-cause dropouts). Secondary outcomes include gadolinium-enhancing lesions, cerebral atrophy and tolerability (dropouts due to adverse events). Neurocognitive measures across three domains including processing speed, working memory and verbal learning will be included as exploratory outcomes. Data will be analysed using a random-effects pairwise meta-analysis and a Bayesian hierarchical random effects network meta-analysis to evaluate the efficacy, acceptability and tolerability of the included DMTs. Subgroup and sensitivity analyses will be conducted to assess the robustness of the findings. The review will be reported using the Preferred Reporting Items for Systematic Reviews incorporating Network Meta-Analyses statement.

Strengths and limitations of this study

► MRI outcomes are frequently used to make treatment decisions around disease-modifying therapies (DMTs) in multiple sclerosis; this will be the first network meta-analysis comparing DMTs in terms of MRI changes as a primary outcome.
► Informed by a comprehensive search strategy, the Bayesian network meta-analysis described in this protocol will synthesise all direct and indirect evidence from multiple treatment comparisons to yield inter-relations of effects across numerous treatments for multiple sclerosis.
► The quality of direct and indirect effect estimates for each outcome will be evaluated using the Grading of Recommendations Assessment, Development and Evaluation framework.
► The limitations of the included studies will be determined using Cochrane’s Risk of Bias tool (RoB 2.0).
► A limitation of this review protocol is the exclusion of articles not published in English, and the exclusion of some DMTs used in jurisdictions other than Australia.

INTRODUCTION

Multiple sclerosis (MS) is a progressive, chronic, autoimmune disease of the central nervous system in which the myelin sheath protecting axons is damaged resulting in distorted nerve signals and pathways.1,2 MS is associated with a complex range of symptoms including visual disturbance, fatigue, pain, reduced mobility, reduced coordination, cognitive impairment and mood changes.
Most patients are initially diagnosed with relapsing-remitting MS (RRMS), characterised by acute clinical attacks (relapses) followed by partial/full recovery and periods of clinical stability (remissions).3 Historically, the management approach has been to tailor therapy based on the efficacy and safety profiles of available agents, disease activity and individual preferences.7 For example, in cases of highly active disease, treatment with higher efficacy DMTs with worse safety profiles may be given early.9 Treatment switching to more potent and higher risk DMTs is also common in cases of ‘breakthrough disease’ (ie, disease activity despite ongoing treatment).8 Several international guidelines have been developed to aid and standardise management decisions in MS. These include the Canadian Multiple Sclerosis Working Group recommendations, the European Committee of Treatment and Research in Multiple Sclerosis and the European Academy of Neurology guideline and the American Academy of Neurology guideline recommendations.9,11 Generally, there are no lines of therapy or thresholds for switching DMTs specified in these guidelines and decisions should be made on a case-by-case basis.12,13

There is an established evidence base that compares DMTs in terms of how they affect relapse rate and disability progression of MS.14–20 However, clinical practice increasingly appears to be aspiring to the goal of no evidence of disease activity (NEDA).6,21–23 Earlier iterations of NEDA, such as NEDA-3, which originally combined the endpoints of relapse, disability progression and T2-weighted or gadolinium-enhanced lesions observed on MRI, were largely based on the sequelae of inflammatory MS activity, which do not consistently correlate with the longer-term neurodegenerative processes which underlie overall disease progression. More recently, parameters for brain volume loss have been added in NEDA-4 to reflect the longer-term aspects of the disease course. Some authors have also proposed adding assessments of neurocognitive function to NEDA assessments.22,24 Neurocognitive impairment, including delayed processing and learning speeds, forms an important component of disability in MS, affecting 40%–65% of patients.25 Neurocognitive impairment has extensive socioeconomic consequences for patients and their families, greatly affecting quality of life.26–28 The impact of cognitive symptoms can be compounded by psychosocial symptoms such as depression, fatigue and anxiety.29 The most commonly recorded cognitive deficits include those that affect memory, attention, verbal fluency, information processing and executive functions.30 Assessment of neurocognitive outcomes in MS is traditionally based on psychometric self-evaluation scales, such as the Symbol Digit Modalities Test.31 Clinically, MRI has an integral role in the diagnosis, management and prognostication of patients with MS.32 MRI is used to assess treatment response, assisting clinicians to decide when to alter doses or switch DMTs. The choice of DMTs is complicated by the increasing array of treatment options with very different risk–benefit profiles. One meta-analysis examining surrogate outcomes for disability progression noted that the average effect of a treatment on MRI outcomes is strongly correlated with its effect on Expanded Disability Status Scale (EDSS) worsening.33 Conversely, another study showed no radiological benefit of switching DMTs for patients with breakthrough MRI activity.34 There are strong correlations between neurocognitive impairment and MRI measures, including whole brain atrophy, third ventricle size and lesion volume.35 A meta-analysis found a robust but small-to-moderate positive effect of DMTs on cognitive test performance in RRMS.36 While the evidence to date is conflicting, it is broadly believed that in RRMS, optimising treatment for MRI outcomes may prevent future relapse and may also lessen or prevent neurocognitive impairment.

To our knowledge, there is only one other network meta-analysis (NMA) that compares the efficacy of these drugs in terms of MRI outcomes (as secondary outcomes).37 While methodologically rigorous, this paper did not include all of the studies that contain data on these outcomes, as these were not the primary outcomes of this study. There is also no NMA that has compared these drugs in terms of neurocognitive outcomes. While these outcomes are being used in practice to make treatment decisions, and indeed are recommended in international treatment guidelines to guide disease management, as yet the performance of DMTs on these outcomes in clinical trials have not been thoroughly simultaneously compared. Further, regulatory and reimbursement bodies continue to make decisions about DMTs primarily based on the outcomes of clinical relapse and disability progression, which does not align with their use in clinical practice. In this study we will perform a systematic review and NMA to examine the relative efficacy and tolerability of 15 DMTs in the long-term treatment of adults with RRMS, in terms of MRI and neurocognitive outcomes.

METHODS AND ANALYSIS

This protocol paper was prepared according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (see online supplemental appendix 1).38 The review commenced in July 2020, and we anticipate the final report will be complete in March 2022.

Eligibility criteria

Types of studies

We will include double-blind Phase II or III randomised controlled trials (RCTs), in which any of the 15 included DMTs (described in the Interventions section) was compared as monotherapy with placebo or another active drug for the treatment of RRMS. Parallel-group RCTs including the first phase of cross-over trials, and multi-arm trials, will be included. Quasi-randomised trials (eg, those
Neurocognitive test performance, including cognitive impairment in MS, will be assessed using standardized neuropsychological tests. We will include trials reporting on patients aged 18 years and over with a diagnosis of RRMS, or those with a patient population composed of ≥80% of patients with RRMS.

Interventions
We will include all DMTs currently reimbursed in Australia for the treatment of RRMS: interferon beta-1a intramuscular, interferon beta-1a subcutaneous, interferon beta-1b, peginterferon beta-1a, glatiramer acetate, natalizumab, ocrelizumab, alemtuzumab, dimethyl fumarate, teriflunomide, fingolimod, cladribine and ozanimod. Further, we will also include rituximab and mitoxantrone, which are used off-label for the treatment of RRMS in Australia. We included these 15 DMTs that have evidence of efficacy for disability progression and relapse, and excluded drugs with unclear or emerging evidence. Although we included daclizumab in the search strategy, we will exclude it from the NMA after it was withdrawn from the market in Australia following reports of encephalitis associated with treatment. We will also exclude combination treatments, interventions involving over-the-counter drugs and all non-pharmacological treatments.

Outcome measures
Primary outcomes
- Number of new or enlarging T1/hypointense lesions.
- Number of new or enlarging T2/hyperintense lesions.
- All-cause dropouts (measure of acceptability), or the proportion of patients who discontinued the trial early due to any reason at the end of the study.

Secondary outcomes
- Number of gadolinium-enhancing lesions.
- Cerebral atrophy.
- Dropouts due to adverse events (measure of tolerability), or the proportion of patients who discontinued the trial early due to adverse events at the end of the study.

Exploratory outcomes
- Neurocognitive test performance, including cognitive processing speed, working memory and verbal learning. Consistent with other meta-analytic reviews investigating cognitive impairment in MS, we will classify any neuropsychological test data into these cognitive domains prior to extraction. The metafor package of the R programming language will be used to compute effect-sizes representing mean differences between treatments for each domain. Table 1 summarises the measures that we expect will be most frequently used to assess the included cognitive domains in the included trials.

Search strategy and data management
Search strategy
We will conduct a systematic search to identify all published, unpublished and ongoing RCTs that compared the efficacy in terms of MRI or neurocognitive outcomes of one drug with another, or with placebo, in the treatment of RRMS. Safety outcomes reported in the included studies, including discontinuations due to any cause and discontinuations due to adverse events, will be used to evaluate the acceptability and tolerability of each DMT. These outcomes have previously been used to compare the safety of included treatments in NMA. The following databases will be searched: Cochrane Central Register of Controlled Trials, Embase, and MEDLINE. We will also search websites of international regulatory bodies for pharmaceuticals (Food and Drug Administration (USA), Medicines and Healthcare Products Regulatory Agency (UK), European Medicines Agency (Europe) and Therapeutic Goods Administration (Australia)) and international trial registries (including ClinicalTrials.gov, Australian and New Zealand Clinical Trials Registry, and WHO International Clinical Trials Registry Platform). The search will be restricted to articles published in English only. In the case of incomplete reporting or unpublished studies, authors will be contacted up to three times to supplement the information in the original papers. We will also hand search the reference lists of included studies for other relevant studies. Full details of the initial draft search strategy are included in the supplementary appendix to this manuscript (online supplemental appendix 2).

Study selection
Each title and abstract retrieved by the search will be independently reviewed by two authors using Covidence software. If both agree that a trial does not meet the eligibility criteria, it will be excluded. Any conflicts will

| Table 1. Measures of neurocognitive performance across the included cognitive domains |
|---------------------------------|---------------------------------------------------------------|
| Cognitive domain               | Most frequently reported neuropsychological test/s            |
| Working/visual memory          | Paced auditory serial addition test 2 s interval (PASAT 2s), PASAT 3 s interval (PASAT 3s), brief visuospatial memory test – revised |
| Processing speed               | Symbol digit modalities test                                 |
| Verbal learning                | California verbal learning test (CVLT) – verbal learning immediate verbal memory, CVLT – immediate recall, Rey auditory visual learning test – immediate recall |

be resolved via consensus with a third author. When consensus cannot be reached that article will be included in the next stage of the review. Following title and abstract screening, full texts of potentially relevant studies will be assessed for eligibility by two authors independently, with any conflicts resolved in consultation with a third member of the review team. Review authors will communicate regularly during the screening process to discuss study selection, with any alterations recorded.

Data extraction
Data extraction and an assessment of risk of bias will be completed independently by two reviewers for each trial. We will use a data extraction template that will be developed for this study. The reliability of data extraction will be checked in at least 20% of included articles by a third author. We will extract information including bibliographic data and trial characteristics (eg, author names, journal, publication year, country, design, blinding and inclusion criteria), participant characteristics (eg, diagnostic criteria for RRMS, age range, duration of disease, prior treatments, severity of disability and average number of previous relapses), intervention details (eg, duration of treatment, dose ranges and mean doses of study drugs) and outcome measures.

Means and SD will be collected for each outcome. If the SD is not reported, we will estimate using p values, t values, CIs or SEs where reported. Where these data are missing, attempts will be made to obtain these data by contacting trial authors. When this fails, they will be borrowed from the other trials in the network or from other published reports.

Trial duration
We will include any trial with an assessment of relevant outcomes at a minimum of 6 months (or 24 weeks). For the primary analysis, we will consider the 12-month endpoint data. We will classify studies into groups per outcome, based on the treatment duration (approximately 6 months, 12 months, and 24 months, depending upon data availability) to assess, using sensitivity analyses, whether the duration of treatment leads to any differences in benefit.

Dosage information
We will include treatment arms in which patients are receiving the licensed dose range for each DMT, with all comparisons between DMTs to be within these ranges. Any trials or treatment arms within trials that used dosages outside of the licensed ranges will be excluded.

Risk of bias assessment
The risk of bias of the included studies will be assessed based on their primary outcome, using the tool described in the Cochrane Collaboration Handbook (RoB 2.0).46 The assessment will be performed by two authors independently, with conflicts resolved by consensus with another author. This includes the following domains: random sequence generation, allocation concealment, blinding of participants, blinding of outcome assessor, incomplete outcome data, selective outcome reporting and other sources of bias, including sponsorship bias. The risk of bias of each study will be classified as low, medium, high, or unclear. Studies will be classified as having overall low risk of bias if none of the domains are rated as high risk of bias, and three or less are rated as unclear; as having overall moderate risk if one domain is rated as high risk of bias, or none are rated as high risk of bias but four or more are rated as unclear risk; and all other studies will be classified as having an overall high risk of bias.

Statistical analysis
Descriptive synthesis of results
We will present a summary of all trial and participant characteristics across all included studies, describing study design and relevant variables.

Pairwise meta-analysis
We will conduct an analysis of all DMTs that were directly compared either head-to-head or with placebo, to investigate the overall effect of DMTs on the treatment of RRMS for all primary and secondary outcomes. We will calculate summary standardised mean differences (SMD, Cohen’s d) for continuous outcomes or ORs for dichotomous outcomes, both with 95% credible intervals (CrI). We will use a random effects meta-analysis model to estimate effect estimates. The heterogeneity parameter for between-study variance will be estimated using restricted maximum likelihood, and the 95% CIs will be derived using the Knapp-Hartung method for all included outcomes. The I^2 statistic and its 95% CrI will be calculated to convey the amount of heterogeneity.

Network meta-analysis
We will use NMA to compare the efficacy and acceptability/tolerability of the various DMTs included in this study with one another, and placebo, for all primary and secondary outcomes. NMA combines direct and indirect evidence for all relative treatment effects and provides estimates with maximum power by allowing more studies to contribute to each treatment comparison. Where different measures are used to assess the same outcome, continuous outcomes data will be pooled with SMD and dichotomous outcomes will be analysed by calculating the OR. We will conduct a random-effects NMA within a Bayesian framework to synthesise all evidence for each outcome. Uninformative prior distributions will be used for all model parameters. Posterior distributions of direct and indirect treatment effect parameters will be simulated using the Markov Chain Monte Carlo method with multiple chains to assess convergence. Posterior distributions will be summarised as means with 95% CrI. We will also estimate the ranking probabilities for all treatments of being at each possible rank for each intervention. Then, we will obtain a treatment hierarchy using the
surface under the cumulative ranking curve and mean ranks.

**Assessment of the transitivity assumption**

A key assumption of NMA is that all included studies are jointly randomisable; that is, that the distribution of potential effect modifiers is balanced across comparisons. There are few clinical features in MS that are known to be effect modifiers; by limiting our samples to participants with RRMS only, we are increasing the likelihood of transitivity in the network. Other variables that may influence the outcomes of treatment efficacy include age, sex, disease severity and duration, healthcare system/setting and switching from previous treatments/treatment duration. These will be tabulated and summarised across treatment comparisons to explore potential imbalances that may result in violations of transitivity.

**Measures for inconsistency in the network**

A key assumption of NMA is that there is consistency in the network, that is, that direct and indirect evidence are in agreement as to the direction of the effect. We will check the assumption of local and global consistency using the nodesplit method using the gemtc R package.

**Publication bias**

Within each pairwise comparison, we will use both the contour-enhanced funnel plot and Egger’s test to assess risk of publication bias. We will also use the comparison-adjusted funnel plots of all trials with placebo controls or inactive controls to investigate whether results in imprecise trials differ from those in more precise trials included in the NMA.

**Subgroup and sensitivity analyses**

We will investigate the source of any significant heterogeneity and/or inconsistency in the networks. Where indicated and appropriate, we will perform subgroup analyses on possible effect modifiers that might have caused the heterogeneity or inconsistency between studies, including age and sex ratio of participants, severity of MS symptoms at baseline, disease duration, previous treatment duration and any treatment switching recorded at baseline. Severity of baseline MS symptoms will be assessed by collecting data on EDSS and annualised relapse rates. We will conduct a number of sensitivity analyses, including investigating the impact of excluding trials with a high risk of bias, and the impact of publication date. Subgroup and sensitivity analyses will investigate both impact on the results of the NMA, and also whether consistency and model fit are affected.

**GRADE quality assessment of a network**

We will assess the quality of evidence contributing to primary outcomes with the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group framework. This characterises the quality of a body of evidence contributing to a NMA based on study limitations, imprecision, heterogeneity or inconsistency, indirectness and publication bias. The quality of evidence will be classified as very low, low, moderate or high.

**Ethics and dissemination**

This NMA does not need ethical approval, as the data to be used are secondary data from published sources. The completed review will be published in a peer-reviewed journal, and findings will be further disseminated through presentation at appropriate forums or conferences.

**Statement of patient and public involvement**

There was no patient and public involvement in the design of this study.

**DISCUSSION**

Network meta-analysis is a novel technique and has previously been used to assess treatment outcomes of DMTs in terms of relapse rate and disability progression in RRMS. This study will be among the first to use network meta-analysis to summarise the evidence for DMTs in the treatment of RRMS in terms of MRI and neurocognitive outcomes. The findings from the analysis will provide hierarchical information on the relative efficacy and acceptability of the DMTs for these outcomes. This will facilitate further understanding for clinicians of the risk–benefit profiles of the various DMTs used in the treatment of RRMS. NMA is a powerful methodology, however it is limited by the availability of relevant data, their quality and their comparability. We expect that data will be available for the proposed MRI outcomes, however there is uncertainty regarding the availability of clinical trials evaluating neurocognitive performance. In the case that we do not identify enough studies to synthesise results for these outcomes, this review will highlight this important gap in the literature.

As the treatment for RRMS is not curative, addressing all possible outcomes that could affect quality of life is imperative when deciding on a management plan. The cognitive consequences of MS have not been studied as extensively as physical changes and disability progression. However, the goal of treatment is changing to include MRI and cognitive treatment targets. As such, the findings of this network meta-analysis will make a significant contribution to this important but under-researched field.

**Twitter** Madeleine Hinwood @madeleinhinwood

**Contributors** Authors MH, WLG, DL and MB contributed to the inception of the concept and rationale for the study design; author AS assisted with the development of a search strategy; authors CO, MH, SL, MD, JEK, VC, MM and NW assisted with the design of data synthesis and analysis. MH, SL, MD, JEK, VC, NW and MM substantially contributed to the writing of this protocol and the work completed, with critical input from WLG, DL and MB. All authors were involved in discussions regarding the progress of the protocol and all were involved in the editing of the manuscript and have read and approved the final version of the protocol.

**Funding** The development of this protocol was supported by the University of Newcastle, Australia (award/grant number is not applicable).
REFERENCES

45 Covidence systematic review software, veritas health innovation, Melbourne, Australia. Available: www.covidence.org