BMC Neurology

RESEARCH

Open Access

Brain volume loss in relapsing multiple sclerosis: indirect treatment comparisons of available disease-modifying therapies

Robert Zivadinov[1,2], Alexander J. Keenan[3*], Hoa H. Le[3], Maria Ait-Tihyaty[4], Kavita Gandhi[4], Matthew L. Zierhut[5], Elizabeth M. Salvo-Halloran[6], Abril Oliva Ramirez[6], Vivian Vuong[6], Sumeet Singh[6] and Brian Hutton[7]

Abstract

Background Brain volume loss (BVL) has been identified as a predictor of disability progression in relapsing multiple sclerosis (RMS). As many available disease-modifying treatments (DMTs) have shown an effect on slowing BVL, this is becoming an emerging clinical endpoint in RMS clinical trials.

Methods In this study, a systematic literature review was conducted to identify BVL results from randomized controlled trials of DMTs in RMS. Indirect treatment comparisons (ITCs) were conducted to estimate the relative efficacy of DMTs on BVL using two approaches: a model-based meta-analysis (MBMA) with adjustment for measurement timepoint and DMT dosage, and a network meta-analysis (NMA).

Results In the MBMA, DMTs associated with significantly reduced BVL versus placebo at two years included fingolimod (mean difference [MD] = 0.25; 95% confidence interval [CI] = 0.15 – 0.36), ozanimod (MD = 0.26; 95% CI = 0.12 – 0.41), teriflunomide (MD = 0.38; 95% CI = 0.20 – 0.55), alemtuzumab (MD = 0.38; 95% CI = 0.10 – 0.67) and ponesimod (MD = 0.71; 95% CI = 0.48 – 0.95), whereas interferons and natalizumab performed the most poorly. The results of NMA analysis were generally comparable with those of the MBMA.

Conclusions Limitations of these analyses included the potential for confounding due to pseudoatrophy, and a lack of long-term clinical data for BVL. Our findings suggest that important differences in BVL may exist between DMTs. Continued investigation of BVL in studies of RMS is important to complement traditional disability endpoints, and to foster a better understanding of the mechanisms by which DMTs can slow BVL.

Keywords Brain volume loss, Model-based meta-analysis, Network meta-analysis, Disease modifying therapy, Systematic literature review, Magnetic resonance imaging, Multiple sclerosis

*Correspondence:
Alexander J. Keenan
AKeenan1@its.jnj.com
[1] Buffalo Neuroimaging Analysis Center, Department of Neurology, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, State University of New York, Buffalo, NY, USA
[2] Center for Biomedical Imaging at the Clinical Translational Science Institute, University at Buffalo, State University of New York, Buffalo, NY, USA
[3] Janssen Scientific Affairs, Janssen Pharmaceuticals, Titusville, NJ, USA
[4] Janssen Pharmaceuticals, Titusville, NJ, USA
[5] Certara USA, Inc., Radnor, PA, USA
[6] EVERSANA, Value & Evidence Services, Burlington, ON, Canada
[7] Ottawa Hospital Research Institute, Ottawa, ON, Canada

Background

Brain volume loss (BVL) detected through magnetic resonance imaging (MRI) can be identified in the earliest stages of multiple sclerosis (MS), and occurs at a faster rate than in healthy adults [1, 2]. De Stefano et al. reported an annual BVL rate of 0.5 to 1.35% for MS patients, compared to a rate of 0.01–0.3% for individuals without the disease [1]. This marker of degeneration has been reliably correlated with future physical and cognitive disability [3, 4], and has been suggested as a key mechanism by which disease-modifying therapies (DMTs) can prevent or delay disability progression [5]. Predictors of disability progression in relapsing multiple sclerosis (RMS) are important to characterize, given that the accumulation of disability increases healthcare costs and greatly diminishes health-related quality of life [6]. Many DMTs have been noted to slow BVL in MS patients [7], potentially through mechanisms that are distinct from the effects of DMTs on central nervous system (CNS) lesions and require further research [8]. Recent evidence indicates that such protective impacts can be sustained over the long-term [9]. As such, BVL is an emerging clinical endpoint that should be considered when evaluating the efficacy of available and emerging DMTs.

While BVL is a commonly evaluated outcome measure in Phase 3 clinical trials in RMS, there is a paucity of head-to-head trials of DMTs. In such cases, indirect treatment comparison (ITC) methods are suitable for comparing treatments, and are often used to inform clinical, regulatory and reimbursement decisions. These methods leverage the existing direct evidence from clinical trials in order to estimate relative treatment effects for interventions that have not been directly compared [10, 11]. Currently, there are several approaches for conducting ITCs. Given the large number of randomized controlled trials (RCTs) reporting BVL for various DMTs, ITC approaches leveraging a network of multiple interventions and underlying RCTs, connected via common comparators, allow for the most comprehensive assessment of relative treatment efficacy. In this study, both model-based meta-analysis (MBMA) and network meta-analysis (NMA) were employed.

NMA is a well-accepted methodology that combines direct and indirect evidence across the network of studies to compare interventions [12–14], and has been extensively presented in the MS literature [15–18]. MBMA similarly utilizes the network of studies to compare interventions. However, MBMA provides a more flexible analytical framework in which additional variables can be readily incorporated in the statistical model [19, 20]. Specifically, MBMA can incorporate data pertaining to various dosages of the same DMT, as well as longitudinal data. This makes MBMA particularly advantageous for an investigation of BVL in RMS, since a majority of large RMS trials have assessed investigational agents at multiple dosages over a range of study durations, with varied measurement timepoints.

The objective of this study was to systematically collect published BVL data from RCTs in RMS, and to compare the effects of various DMTs on BVL using MBMA and NMA.

Methods

Systematic literature review methods

The systematic literature review (SLR) was reported in accordance with PRISMA reporting guidelines [21]. Searches of EMBASE, and MEDLINE®, and the Cochrane Central Register of Controlled Trials were designed by an information specialist and peer reviewed by a second information specialist [22]. Search strategies used a combination of controlled vocabulary (e.g., "multiple sclerosis") and keywords (e.g., "relapsing remitting"), with no date restrictions. The database searches were conducted in August, 2021 (search strategies are provided in Online Resource 1.1). Searches of several additional sources were also conducted, including the reference lists of included primary publications and relevant systematic reviews, as well as proceedings of large multiple sclerosis congresses, clinical trial registries (e.g., ClinicalTrials.gov), and regulatory and health technology assessment (HTA) agency websites. These additional sources were searched in late 2021; further details including date restrictions are provided in Online Resource 1.2.

The eligibility criteria for the SLR are detailed in Online Resource 1.3. English language citations reporting data from RCTs comparing two or more regimens of interest (e.g., placebo, interferon DMTs, orally administered DMTs, or monoclonal antibody DMTs) on MRI outcomes in adult RMS populations were eligible for inclusion. RMS studies were eligible for inclusion if the proportion of patients with relapsing–remitting multiple sclerosis (RRMS) was ≥ 80%. Two reviewers independently screened citations using Distiller SR (Evidence Partners; Ottawa, Canada), first based on titles and abstracts, then based on full text review. A single reviewer collected data regarding study populations, treatment regimens, and MRI outcomes using a structured data extraction form in Microsoft Excel (Microsoft, Seattle, Washington, USA), which was validated by a second reviewer. Results for BVL collected using any technique (e.g., Structural Image Evaluation using Normalization of Atrophy [SIENA], or brain parenchymal fraction calculation [BPF]), where reported as a mean change from baseline per treatment group, were abstracted from included studies. Otherwise, median change from baseline in BVL was collected, or

data which lacked such descriptors. Variance measures were also collected for each estimate, where available (e.g., standard deviation [SD], standard error [SE], 95% confidence interval [CI]). The Cochrane Risk of Bias tool was used to appraise the quality of included RCTs [23], performed by one reviewer and validated by a second reviewer. Disagreements between reviewers pertaining to study selection, data collection, and quality assessment were resolved by discussion, or by a third reviewer.

ITC feasibility assessment

The validity of ITCs relies on whether there are systematic differences among the included studies, especially in patient characteristics that are treatment effect modifiers [24–27]. Failure to account for these differences can lead to biased comparisons of treatment effect [28, 29]. The feasibility of conducting ITCs was assessed qualitatively by comparing key study and patient characteristics across trials included in the SLR that reported BVL data. Assessment of cross-trial differences was conducted for various study design elements (e.g., trial phase, blinding, sample size) using summary tables, and for patient characteristics (e.g., age, gender, baseline Expanded Disability Status Scale [EDSS], prior DMT exposure) using bar charts.

The measurement timepoints and outcome definitions were also compared between studies. Any measurement techniques for BVL were considered eligible for inclusion. The use of different measurement techniques was not expected to introduce bias to relative effect estimates. This assumption is aligned with published meta-analyses of BVL [7, 30]. Notably, the relevance of BVL data measured at or prior to 24 weeks of DMT treatment was considered limited, due to potential confounding caused by pseudoatrophy. As such, only BVL measurements collected beyond 24 weeks were considered eligible for inclusion in ITCs.

Selection of study data for ITCs

Mean change-from-baseline data was preferred for incorporation in ITCs, however median changes were considered if mean estimates were not reported. Reported measures of variance were either converted to SD, or where no such measure was available, SD values were imputed using the following equation:

$$SD(\%BVL)_{ij} = \beta_0 + \beta_1 \cdot |\%BVL|_{ij} \quad (1)$$

Where β_0 and β_1 were estimated based on the linear relationship between BVL and its SD, weighted by sample size, in the j^{th} arm of the i^{th} trial.

MBMA methods

MBMA methods aligned with practices published by Mandema et al. [31, 32]. All MBMAs were conducted using the R programming language version 3.6.3 (R Core Team, Vienna, Austria) with the nlme() library version 3.1–144. All data points after 24 weeks and all tested doses of each included drug were utilized in the MBMA.

As multiple time points and multiple doses were included in the MBMA, a longitudinal, dose–response model was utilized to explore the influence of time and dose on relative treatment effect (i.e., BVL relative to placebo). Both fixed effect and random effect approaches were explored to determine the most appropriate model based on an ANOVA test. A dose response power function was explored separately for all DMTs with data at multiple doses. Time was incorporated as a covariate on relative effects using a power function anchored to 104 weeks: $[\text{time}/104]^k$. Additionally, baseline mean number of T1 gadolinium (Gd+) enhancing lesions was explored as an influential covariate on relative treatment effect. For this specific investigation, a subset of trials that did not report any relevant T1 Gd+lesion information were excluded.

Additional MBMA methodology details (including a description of the model with equation) are provided in Online Resource 2.1.

NMA methods

Bayesian NMAs were performed using the R programming language version 3.5.3 and Just Another Gibbs Sampler (JAGS) version 4.3.0. NMA models were based on code adapted from the National Institute of Health and Care Excellence (NICE) Evidence Synthesis Decision Support Unit (DSU) Technical Support Document (TSD) Series [33]. Only DMT dosages that are currently authorized for use in the US were included. Given the progressive decline of BVL over time, only data collected at approximately the same timepoint, i.e., two years, were included in the NMA. This was aligned with the MBMA, which used two years (104 weeks) as the reference timepoint for predicted BVL. Additional details regarding the NMA methods, model fit, and inconsistency analysis are provided in Online Resource 2.2.

Results interpretation for indirect treatment comparisons

For both MBMA and NMA, effect estimates reported within Forest plots reflected the probability of a DMT regimen being better than placebo, based on being to the right or left of the Y axis. Where the interval pertaining to an effect estimate did not cross the Y axis, this indicated significant outperformance of the DMT regimen over placebo. The relative distances of effect estimates

from the Y axis also indicated the relative rank of DMT regimens. For the NMA, a league table reporting all possible pairwise comparisons is provided in the Online Resource 5.0.

Results

Evidence identified in the systematic literature review

The SLR identified 40 RCTs, across 158 individual reports (Online Resource 3.1 and 3.2). More than 30% of these reports were congress proceedings, trial registry records, and HTA or regulatory agency documents. In total, 31 trials reported data on BVL. Quality assessments broadly indicated low risks of bias for these trials (Online Resource 3.3). A summary of the main characteristics of these trials is provided in Table 1, including details regarding the reporting of BVL.

Table 1 considers all RCTs reporting BVL in any form, for the DMTs in the scope of this review. The Phase 2 trials of ponesimod and ocrelizumab were limited to 24 weeks in duration [34, 35]. Similarly, the only BVL results reported from baseline in the CLARITY trial were collected at 24 weeks [36]. Following the exclusion of these three trials, 28 trials reporting BVL between 36 and 156 weeks were considered eligible for ITCs. Missing standard deviation values were imputed as described above (Eq. 1), with estimated parameters: $SD(\%BVL)_{ij} = 0.558 + 0.713 \cdot |\%BVL|_{ij}$.

Feasibility of conducting indirect treatment comparisons

The ITC eligible trials were found to be comparable with regards to study design, patient characteristics, and measurement/reporting of total BVL (see Online Resource 4.1 through 4.4 for supporting tables and plots). The majority of studies were multi-national, phase III, double-blinded, and included follow-up of approximately two years.

Moreover, the measurement technique used to collect BVL varied across trials; most trials indicated the use of either SIENA or BPF calculation, with SIENA being more frequently utilized in recent trials.

Indirect treatment comparison findings

The full network of evidence included in the ITCs is provided in Fig. 1. All trial data leveraged for the ITCs is provided in Online Resource 5.0. For the MBMA, the investigation of dose–response considered all drugs for which multiple doses were included: alemtuzumab, dimethyl fumarate, fingolimod, interferon β-1b, ozanimod, peginterferon β-1a, and teriflunomide. No significant difference was found between dimethyl fumarate 240 mg (twice daily) and dimethyl fumarate 240 mg (three times daily), thus these doses were pooled, and a single treatment effect was estimated for dimethyl fumarate. Peginterferon β-1a and interferon β-1b had a negligible dose–response relationship ($p > 0.75$), and a pooled effect estimate was incorporated for each respective drug. Although not statistically significant ($p > 0.05$), a dose–response relationship was incorporated for the remaining four drugs with data at multiple doses (alemtuzumab, fingolimod, ozanimod, teriflunomide) to ensure that any potential dose-related variability was accounted for.

Upon incorporation of time as a covariate on relative effects, results demonstrated that a relative treatment effect at 52 weeks would be approximately 38% smaller than the same treatment effect at 104 weeks [95% CI: 17%—53%] based on the time exponent (k) estimation of 0.70 with 95% CI [0.27 – 1.1]. Compared to a model excluding any influence of time on relative effect, the final MBMA described the data significantly better ($p < 0.01$ from ANOVA test).

A model adjusting for the number of T1 gadolinium enhancing lesions (T1 Gd+) at baseline was also considered. However, this model was not significantly better fitting than the final model, which excluded the influence of baseline T1 Gd+ lesions ($p > 0.45$).

A random effects model was also explored. Using an ANOVA test, there was a negligible difference in how well this model described the data versus the fixed effects model.

Considering the final MBMA incorporating adjustment for dosage and timepoint, the DMTs predicted to significantly outperform placebo on BVL in order of decreasing relative efficacy included: ponesimod, alemtuzumab, teriflunomide, ozanimod, and fingolimod (Fig. 2a). That is, statistically significant benefits versus placebo were noted in each case. Interferons and natalizumab appeared to perform the most poorly in this analysis. A league table reporting the full results of the MBMA can be found in Online Resource 6.0.

Although the confirmatory NMA utilized data from the same clinical trials as the MBMA, only data collected at approximately two years, for treatment dosages that are currently authorized for use in the US, were included (Fig. 1).

The effect estimates from the confirmatory NMA were generally associated with greater uncertainty (i.e., wider 95% credible intervals) as compared to the MBMA (Fig. 2b). However, the main findings were consistent. In order of decreasing relative efficacy, the DMTs predicted to significantly outperform placebo included: ponesimod, fingolimod, teriflunomide, and ofatumumab. That is, statistically significant benefits versus placebo were noted in each case. The NMA results broadly aligned with the MBMA results in that most DMT regimens, namely the first-generation injectable therapies, were not predicted to significantly reduce BVL compared with placebo. The

Table 1 Characteristics of RCTs reporting BVL identified in the SLR

Trial name (ClinicalTrials.gov ID)	Study characteristics				Regimens		BVL outcome	
	Phase	*Blinding*	*Enrolment period*	*Geographic location*	*Intervention*	*Comparator*	*Timepoints assessed (from baseline)*	*Measurement technique*
ADVANCE (NCT00906399)	3	Double-blind	2009–2011	NA, EUR, other	• peginterferon 125 ug Q2W • peginterferon 125 ug Q4W	placebo	48 weeks, 96 weeks	Unclear
AFFIRM (NCT00027300)	3	Double-blind	2001-NR	NA, EUR, AU/NZ	natalizumab 300 mg IV Q4W	placebo	52 weeks, 104 weeks	BPF
ASCLEPIOS I (NCT02792218)	3	Double-blind	2016–2018	NA, SA, EUR, AU, RUS	ofatumumab 20 mg Q4W	teriflunomide 14 mg qd	52 weeks, 104 weeks	Unclear
ASCLEPIOS II (NCT02792231)	3	Double-blind	2016–2018	NA, SA, EUR, AU, RUS	ofatumumab 20 mg Q4W	teriflunomide 14 mg qd	52 weeks, 104 weeks	Unclear
ASSESS (NCT00340834)	3b	Double-blind	2012-NR	NA, SA	• fingolimod 0.25 mg qd • fingolimod 0.5 mg qd	glatiramer acetate 20 mg qd	52 weeks	Unclear
BEYOND (NCT00099502)	3	Double-blind	2003–2005	NA, SA, EUR, AU, RUS	• interferon beta-1b 250 ug qod · interferon beta-1b 500 ug qod	glatiramer acetate 20 mg qd	110 weeks	SIENA
BRAVO (NCT00605215)	3	Double-blind	2008-NR	NA, EUR, RUS, AF	*No eligible regimens*[b]	• placebo • interferon beta-1a 30 ug intramuscular qw	104 weeks	SIENA
CAMMS223 (NCT00050778)	2	Single-blind	2002–2004	NA, EUR	• alemtuzumab 12 mg qd × 5 in month 1, qd × 3 in month 12, qd × 3 in month 24 at discretion • alemtuzumab 24 mg qd × 5 in month 1, qd × 3 in month 12, qd × 3 in month 24 at discretion	interferon beta-1a 44 µg subcutaneous tiw	156 weeks	SBV
CARE-MS I (NCT00530348)	3	Single-blind	2007–2009	NA, SA, EUR, AU, RUS	alemtuzumab 12 mg qd × 5 in month 1, qd × 3 in month 12	interferon beta-1a 44 µg subcutaneous tiw	104 weeks	BPF
CARE-MS II (NCT00548405)	3	Single-blind	2007–2009	NA, SA, EUR, AU, RUS	• alemtuzumab 12 mg qd × 5 in month 1, qd × 3 in month 12 • alemtuzumab 24 mg qd × 5 in month 1, qd × 3 in month 12	interferon beta-1a 44 µg subcutaneous tiw	104 weeks	BPF
CLARITY (NCT00213135)[a]	3	Double-blind	2005–2007	NA, SA, EUR, AU, RUS, AF	• cladribine 3.5 mg/kg • cladribine 5.25 mg/kg	placebo	24 weeks, 96 weeks[c]	SIENA

Table 1 (continued)

Trial name (ClinicalTrials.gov ID)	Study characteristics				Regimens		BVL outcome	
	Phase	*Blinding*	*Enrolment period*	*Geographic location*	*Intervention*	*Comparator*	*Timepoints assessed (from baseline)*	*Measurement technique*
CONFIRM (NCT00451451)	3	Double-blind	2007-NR	NA, EUR, NZ	• dimethyl fumarate 240 mg bid • dimethyl fumarate 240 mg tid	• placebo • glatiramer acetate 20 mg qd	48 weeks, 96 weeks	SIENA
COPOLYMER I (NCT00004814)	3	Single-blind	1991-NR	NA (USA only)	glatiramer acetate 20 mg qd	placebo	52 weeks	BPF
DEFINE (NCT00420212)	3	Double-blind	2007-NR	NA, EUR, AU/NZ, AF	• dimethyl fumarate 240 mg bid • dimethyl fumarate 240 mg tid	placebo	48 weeks, 96 weeks	SIENA
European/Canadian GA (n/a)	3	Double-blind	1997–1997	NA (Canada only), EUR	glatiramer acetate 20 mg qd	Placebo	36 weeks	CBV
FREEDOMS (NCT00289978)	3	Double-blind	2006–2007	NA, EUR, AUS, RUS, AF	• fingolimod 0.5 mg qd • fingolimod 1.25 mg qd	Placebo	52 weeks, 104 weeks	SIENA
FREEDOMS II (NCT00355134)	3	Double-blind	2006–2009	NA, EUR, AUS	• fingolimod 0.5 mg qd • fingolimod 1.25 mg qd	Placebo	52 weeks, 104 weeks	SIENA
GALA (NCT01067521)	3	Double-blind	2010-NR	NA, EUR, RUS, AF	glatiramer acetate 40 mg tiw	Placebo	52 weeks	SIENA
GATE (NCT01489254)	3	Double-blind	2011–2013	NA, EUR, RUS, AF	• glatiramer acetate 20 mg qd (brand name) • glatiramer acetate 20 mg qd (generic)	Placebo	39 weeks	Unclear
GOLDEN (NCT01333501)	4	Open-label	2011-NR	EUR	fingolimod 0.5 mg qd	interferon beta-1b 250 µg qod	78 weeks	SIENA
MSCRG (n/a)	3	Double-blind	1990-NR	NA (USA only)	interferon beta-1a 30 ug intramuscular qw	Placebo	52 weeks, 104 weeks	BPF
OPERA I (NCT01247324)	3	Double-blind	2011–2013	NA, SA, EUR, AU, RUS, AF	ocrelizumab 600 mg Q24W	interferon beta-1a 44 µg subcutaneous tiw	96 weeks	SIENA
OPERA II (NCT01412333)	3	Double-blind	2011–2013	NA, SA, EUR, RUS	ocrelizumab 600 mg Q24W	interferon beta-1a 44 µg subcutaneous tiw	96 weeks	SIENA
OPTIMUM (NCT02425644)	3	Double-blind	2015–2019	NA, EUR, RUS	ponesimod 20 mg qd	teriflunomide 14 mg qd	60 weeks, 108 weeks	SIENA
Phase 2 ocrelizumab trial (NCT00676715)[a]	2	Double-blind	2008-NR	NA, EUR, RUS	• ocrelizumab 600 mg Q24W • ocrelizumab 2000 mg Q24W	• interferon beta-1a 30 µg intramuscular qw • placebo	24 weeks	BPF
Phase 2 ponesimod trial (NCT01006265)[a]	2	Double-blind	2009–2010	NA, EUR, AUS, RUS	• ponesimod 10 mg qd • ponesimod 20 mg qd • ponesimod 40 mg qd	Placebo	24 weeks	SIENA

Table 1 (continued)

Trial name (ClinicalTrials.gov ID)	Study characteristics				Regimens		BVL outcome	
	Phase	*Blinding*	*Enrolment period*	*Geographic location*	*Intervention*	*Comparator*	*Timepoints assessed (from baseline)*	*Measurement technique*
RADIANCE B (NCT02047734)	3	Double-blind	2013–2015	NA, EUR, RUS, AF	• ozanimod 0.5 mg qd • ozanimod 1 mg qd	interferon beta-1a 30 µg intramuscular qw	52 weeks, 104 weeks	SIENA
REGARD (NCT00078338)	4	Open-label	2004–2004	NA, SA, EUR, RUS	glatiramer acetate 20 mg qd	interferon beta-1a 44 µg subcutaneous tiw	48 weeks, 96 weeks	SIENA
SUNBEAM (NCT02294058)	3	Double-blind	2014–2015	NA, EUR, NZ, RUS	• ozanimod 0.5 mg qd • ozanimod 1 mg qd	interferon beta-1a 30 µg intramuscular qw	52 weeks	SIENA
TEMSO (NCT00134563)	3	Double-blind	2004–2008	NA, SA, EUR, RUS	• teriflunomide 7 mg qd • teriflunomide 14 mg qd	Placebo	52 weeks, 104 weeks	SIENA
TRANSFORMS (NCT00340834)	3	Double-blind	2006–2007	NA, SA, EUR, AUS, AS	• fingolimod 0.5 mg qd • fingolimod 1.25 mg qd	interferon beta-1a 30 µg intramuscular qw	52 weeks	SIENA

Abbreviations: *AF* Africa, *AS* Asia, *AU* Australia, *bid* twice daily, *BPF* brain parenchymal fraction, *BVL* brain volume loss, *CBV* central brain volume, *EUR* Europe, *NA* North America, *NZ* New Zealand, *Q#W* every # weeks, *qd* once daily, *qod* every other day, *qw* once weekly, *RUS* Russia, *SA* South America, *SBV* supratentorial brain volume, *SIENA* Structural Image Evaluation using Normalisation of Atrophy, *tid* three times daily, *tiw* three times per week

[a] RCTs were included in the SLR, but deemed ineligible for inclusion in ITCs

[b] The main intervention assessed in the BRAVO trial was laquinimod (0.6 mg orally, once daily). This intervention is excluded herein, given that the prespecified eligibility criteria for the SLR did not include laquinimod as a regimen of interest

[c] The latest timepoint of BVL measurement in the CLARITY trial was 96 weeks, however the data did not reflect change from baseline. The other reported timepoint of BVL measurement was 24 weeks, which was ineligible for inclusion in the ITCs (see Methods). Therefore, none of the BVL data from this trial was included in ITCs

[d] The latest timepoint of BVL measurement in the Phase 2 trials of ocrelizumab and ponesimod was 24 weeks, which was ineligible for inclusion in the ITCs (see Methods), therefore the data from these trials were not included in ITCs

effect estimate for ofatumumab was significant in the NMA but not the MBMA, whereas the effect estimates for ozanimod and alemtuzumab were significant in the MBMA but not the NMA. A league table reporting the full results of the NMA can be found in Online Resource 7.0.

Discussion

To our knowledge, this study represents the first attempt to use ITC methods for the purpose of comparing BVL outcomes across various DMTs in patients with RMS. The MBMA was the primary approach for ITCs, as it allowed for adjustment of key variables such as dosage and timepoint, whereas the NMA was chosen to serve as a confirmatory analysis given its wide use and familiarity among clinicians and health care decision-makers. Across both analyses, the DMTs which significantly outperformed placebo included fingolimod, teriflunomide, and ponesimod, where the latter always ranked first.

BVL is not currently considered a core outcome in clinical trials of RMS, although it has become more commonly investigated in recent years. Traditionally, lesion burden assessed through MRI has been used as a surrogate marker for disease activity in RMS [37]; however, the link between lesion burden and other clinical findings has been weak to modest [38]. For example, although the appearance of T2 hyperintense lesions is correlated with disability accumulation in MS, this relationship plateaus at higher disability levels [39]. Nevertheless, it is well understood that available DMTs reduce new lesion activity in RMS, and lesion activity is routinely evaluated in clinical trials. In contrast to lesion counts, BVL is a measure of whole-brain atrophy. Notably, both lesion load and atrophy (including that of white matter and grey matter) have been correlated with disability status in MS. A study comparing MS patients and controls conducted by Tedeschi et al. concluded that amongst lesion load, white matter atrophy, and grey matter atrophy, the latter is the most significant MRI variable in determining the final disability level of MS patients [40]. Ghione et al. used a similar study design to ascertain a correlation between BVL and disability progression [4]. Furthermore, Sprenger et al. have highlighted the potential predictive value of BVL earlier in the disease course on long-term disability

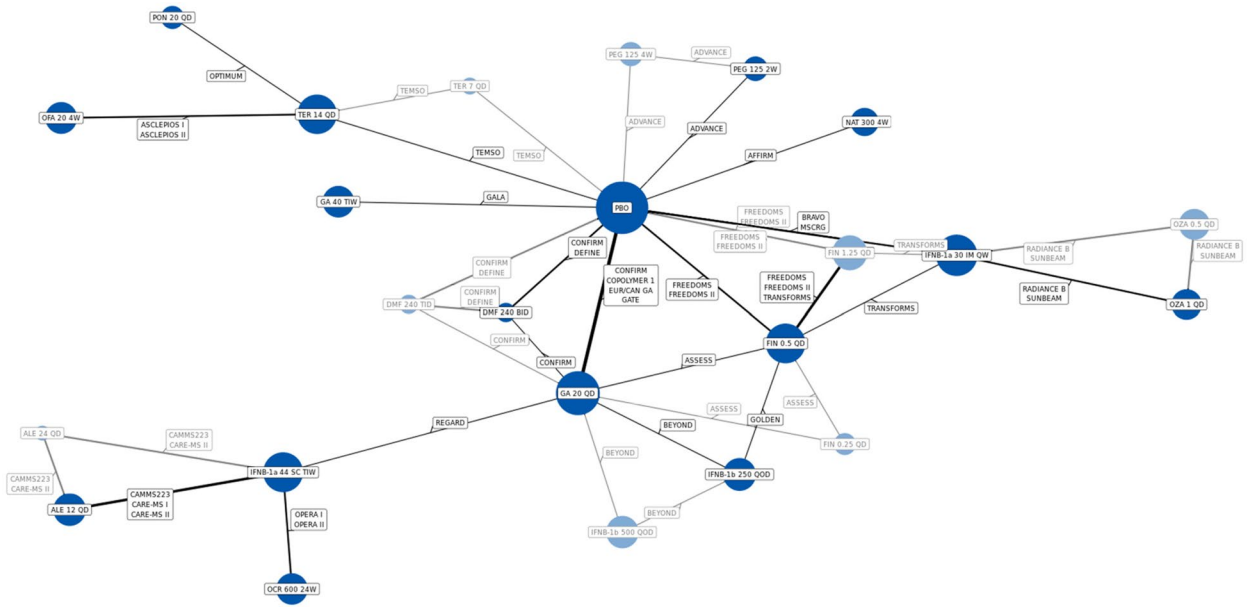

Fig. 1 Evidence network for ITCs of BVL[a]

[a]Treatment nodes are sized proportionally to sample sizes, and connections between treatment nodes are indicated with line thickness proportional to the number of trials informing the connection. All connections were incorporated in the MBMA analysis, whereas only the nodes and trials shown in boldface were included in the NMA (since DMT dosages that were not of interest were excluded from the NMA)

Abbreviations: 24W = every 24 weeks, 2W = every two weeks, 4W = every four weeks, ALE = alemtuzumab, BID = twice daily, BVL = brain volume loss, CI = confidence interval, DMF = dimethyl fumarate, FIN = fingolimod, GA = glatiramer acetate, IFNβ = interferon β, NAT = natalizumab, OFA = ofatumumab, OZA = ozanimod, PBO = placebo, PEG = peginterferon β-1a, PON = ponesimod, Q2W = every two weeks, QD = once daily, QOD = every other day, QW = once weekly, TER = teriflunomide, TID = three times daily

outcomes based on clinical data from the TEMSO extension study [5]. Finally, Sormani et al. have described the prediction of disability accumulation by brain atrophy through a meta-analysis of RCTs.(3) Given the correlation between BVL and disability, a greater understanding of how various treatments may impact BVL is likely to be valuable to clinicians and decision makers.

It is notable that of the three sphingosine-1-phosphate receptor subtype 1 (S1P1) modulators included in our analyses, ponesimod and fingolimod were found to significantly outperform placebo across analyses. Additional research is warranted to understand if a mechanistic link exists between S1P1 regulation and the rate of BVL in RMS. Ozanimod was found to significantly outperform placebo in the MBMA, but the same comparison did not reach statistical significance in the NMA.

In general, monoclonal antibody DMTs are noted for their efficacy on disability outcomes such as 3-month and 6-month disability progression, as demonstrated in published ITCs [41]. However, monoclonal antibody DMTs did not appear to outperform most other agents in the ITCs of BVL. Natalizumab in particular performed poorly overall in comparison to several DMT comparators. Compared with the MBMA results, the effect estimates for the monoclonal antibodies shifted slightly in the NMA, but overall conclusions were similar. Statistically significant benefits versus placebo were not consistently noted between analyses for alemtuzumab, ofatumumab or ocrelizumab.

Our study has several strengths. Broadly, the SLR conforms to guidance from the Cochrane Collaboration and PRISMA reporting guidelines. The search strategy was comprehensive in that it included several database and grey literature sources and was peer-reviewed by a second medical information librarian in advance of conducting searches. ITCs were informed by RCTs identified in the systematic literature review, which were deemed to be of high quality overall using the Cochrane Risk of Bias tool. The methodology for conducting ITCs was aligned with best practices outlined by major HTA agencies such as NICE (NMA), or found in the published literature (MBMA). Notably, the MBMA is able to simultaneously account for multiple potential sources of variability through adjustment, and incorporate data reported at different timepoints and for different dosages. Hence, the current MBMA allows for incorporation of all trial data

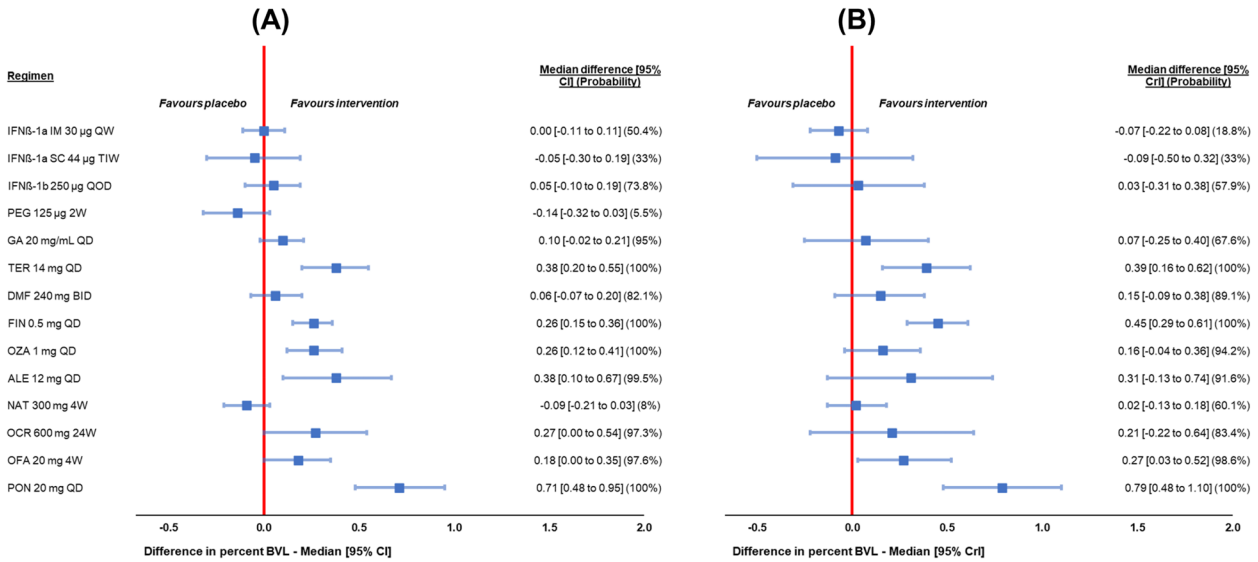

Fig. 2 MBMA (**A**) and NMA (**B**) results for differences in brain volume loss at two years, versus placebo[a]

[a]MBMA and NMA used fixed effect models. Measurement timepoint and dosage were covariates in the MBMA. Round parentheses indicate the probability of being better than placebo. Peginterferon could not be incorporated in the NMA due to a lack of eligible BVL data reported at two years

Abbreviations: 24W = every 24 weeks, 2W = every two weeks, 4W = every four weeks, ALE = alemtuzumab, BID = twice daily, BVL = brain volume loss, CI = confidence interval, DMF = dimethyl fumarate, FIN = fingolimod, GA = glatiramer acetate, IFNβ = interferon β, NAT = natalizumab, OFA = ofatumumab, OZA = ozanimod, PEG = peginterferon β-1a, PON = ponesimod, Q2W = every two weeks, QD = once daily, QOD = every other day, QW = once weekly, TER = teriflunomide

regardless of dosage or timepoint, whereas some data was necessarily excluded from our NMA (e.g., timepoints other than approximately two years) to ensure that effect estimates were sufficiently comparable to yield a valid analysis. To ensure the general alignment of these strategies, the MBMA was contrasted against an NMA which leveraged a more limited data set.

Some limitations of this study should also be considered. Heterogeneity across studies may have had unanticipated impacts on the validity of ITCs. Though patient populations were considered adequately similar for incorporation in ITCs based on the feasibility assessment, there was substantial variation in prior DMT use across trials. If this trait (or other unmeasured traits) are effect modifiers, then variation therein threatens the validity of presented ITCs. Unfortunately, network meta-regression was not feasible to assess the potential impact of prior DMT variation due to the number of single-trial connections in the network, an issue that has been noted in previous ITCs of MS therapies [42]. Future studies could consider the use of more detailed techniques such as propensity score matching or matching-adjusted indirect comparisons to correct for cross-trial differences, where data availability permits (eg, thorough reporting of patient baseline characteristics and access to individual patient data).

Differences over time were also noted in regards to the techniques used for BVL measurement, with the SIENA technique being more predominant in the most recently conducted studies. However, the heterogeneity in measurement techniques between studies was not considered to threaten the validity of ITCs, given that they rely on relative differences between treatment arms per trial, which are expected to be less prone to confounding due to measurement technique than absolute changes in a single trial arm. Fisher et al. reported that BPF and SIENA methods of BVL measurement have robust correlation at baseline and for brain volume change in MS patients, though different absolute results are generated [43]. This high degree of correlation supports the validity of including both measures in NMAs of BVL. The phenomenon of pseudoatrophy was also considered to potentially impact the validity of these analyses. While it is generally understood that pseudoatrophy mostly occurs in the first few months after DMT initiation, there is uncertainty regarding the specific timepoint beyond which the impacts of pseudoatrophy might be negligible [44–46]. Transient volume changes could also result in confounding of BVL measurements, for example, if pseudoatrophy is more likely with one drug class versus another, however limited data are available to inform this issue. In an effort to

reduce the potential for bias due to pseudoatrophy, we excluded BVL measurements collected at or before 24 weeks from the start of DMT treatment. Further research is needed to fully understand this phenomenon, and the extent to which it may confound BVL data collected for various DMT regimens. Finally, it should be noted that additional clinical data is needed to ascertain the longer-term impacts of BVL prevention, as most data points identified as eligible for incorporation in this study were collected at or before two years of treatment.

Conclusion

The MBMA indicated that S1P1 agents fingolimod, ponesimod and ozanimod, as well as teriflunomide and alemtuzumab significantly outperformed placebo on the BVL outcome, after adjustment for measurement timepoint and DMT dosage. These results were confirmed using NMA, which yielded results that were broadly aligned. Some limitations of these analyses included the potential for confounding introduced by pseudoatrophy, and a lack of long-term clinical data. Nevertheless, the differentiation of various DMTs substantiates the importance of evaluating BVL as a clinical endpoint in future trials, as a complement to traditional measures of disability. Further investigations regarding the mechanisms by which various DMTs may reduce BVL rate are also warranted, as are studies that improve our understanding of the long-term relationship between BVL rate and disability progression.

Abbreviations

BPF	Brain parenchymal fraction calculation
BVL	Brain volume loss
CI	Confidence interval
CNS	Central nervous system
DMTs	Disease-modifying treatments
DSU	Decision Support Unit
EDSS	Expanded Disability Status Scale
Gd +	Gadolinium
HTA	Health technology assessment
ITCs	Indirect treatment comparisons
JAGS	Just Another Gibbs Sampler
MBMA	Model-based meta-analysis
MD	Mean difference
MRI	Magnetic resonance imaging
MS	Multiple sclerosis
NICE	National Institute of Health and Care Excellence
NMA	Network meta-analysis
RCTs	Randomized controlled trials
RMS	Relapsing multiple sclerosis
RRMS	Relapsing–remitting multiple sclerosis
S1P1	Sphingosine-1-phosphate receptor subtype 1
SD	Standard deviation
SE	Standard error
SIENA	Structural Image Evaluation using Normalization of Atrophy
SLR	Systematic literature review
T1 Gd +	T1 gadolinium enhancing lesions
TSD	Technical Support Document

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12883-024-03888-6.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements

The authors would like to thank Sandra Vigelis, Alexis Jenkins, and Sarah Da Silva for their assistance in the conduct of the literature review and indirect treatment comparisons.

Authors' contributions

ESH, VV, and AOR conducted the systematic literature review and network meta-analyses. MLZ conducted model-based meta-analyses. AK, HL, KG, MAT, SS, and BH assisted with the interpretation of the data and critically reviewed for importance of intellectual content. RZ provided clinical expertise and reviewed for importance of intellectual content. All authors reviewed and approved the final version of the manuscript.

Funding

This manuscript received funding from Janssen Pharmaceuticals.

Availability of data and materials

All data generated or analysed during this study are included in this published article [and its supplementary information files].

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

ESH, VV, AOR and SS are employees of EVERSANA, and MLZ is an employee of Certara USA. AK, HL, KG and MAT are employees of Janssen, part of Johnson and Johnson group of companies and may hold stock/stock options of Johnson and Johnson. BH has previously received honoraria from EVERSANA for methodologic guidance related to the conduct of systematic reviews and meta-analysis. RZ has received personal compensation from Bristol Myers Squibb, EMD Serono, Sanofi, Keystone Heart, Janssen, 415 Capital, Mapi Pharma, Protembis and Novartis for speaking and consultant fees. He received financial support for research activities from Sanofi, Novartis, Bristol Myers Squibb, Octave, Mapi Pharma, Keystone Heart, Protembis and V-WAVE Medical. The preparation of this manuscript was funded by Janssen.

Received: 14 November 2023 Accepted: 27 September 2024
Published online: 08 October 2024

References

1. De Stefano N, Stromillo ML, Giorgio A, Bartolozzi ML, Battaglini M, Baldini M, et al. Establishing pathological cut-offs of brain atrophy rates in multiple sclerosis. J Neurol Neurosurg Psychiatry. 2016;87(1):93.
2. De Stefano N, Giorgio A, Battaglini M, Rovaris M, Sormani MP, Barkhof F, et al. Assessing brain atrophy rates in a large population of untreated multiple sclerosis subtypes. Neurology. 2010;74(23):1868–76.
3. Sormani MP, Arnold D, De Stefano N. Combined MRI measure of active lesions and brain atrophy as a surrogate for disability in multiple sclerosis: a meta-analysis of randomized trials (P07.096). Neurology. 2016;80(7 Supplement):P07.096.

4. Ghione E, Bergsland N, Dwyer MG, Hagemeier J, Jakimovski D, Paunkoski I, et al. Brain atrophy is associated with disability progression in patients with MS followed in a clinical routine. Am J of Neuroradiol. 2018;39(12):2237–42.
5. Sprenger T, Kappos L, Radue EW, Gaetano L, Mueller-Lenke N, Wuerfel J, et al. Association of brain volume loss and long-term disability outcomes in patients with multiple sclerosis treated with teriflunomide. Mult Scler. 2020;26(10):1207–16.
6. Jones E, Pike J, Marshall T, Ye X. Quantifying the relationship between increased disability and health care resource utilization, quality of life, work productivity, health care costs in patients with multiple sclerosis in the US. BMC Health Serv Res. 2016;16(1):294.
7. Branger P, Parienti J-J, Sormani MP, Defer G. The effect of disease-modifying drugs on brain atrophy in relapsing-remitting multiple sclerosis: a meta-analysis. PLoS ONE. 2016;11(3):e0149685.
8. Wang C, Barnett MH, Yiannikas C, Barton J, Parratt J, You Y, et al. Lesion activity and chronic demyelination are the major determinants of brain atrophy in MS. Neurol Neuroimmunol Neuroinflamm. 2019;6(5):e593.
9. Temmerman J, Van Der Veken F, Engelborghs S, Guldolf K, Nagels G, Smeets D, et al. Brain volume loss can occur at the rate of normal aging in patients with multiple sclerosis who are free from disease activity. J Clin Med. 2022;11(3):523.
10. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol. 1997;50(6):683–91.
11. Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. Methods for population-adjusted indirect comparisons in health technology appraisal. Med Decis Making. 2018;38(2):200–11.
12. Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. Res Synth Methods. 2012;3(2):80–97.
13. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med. 2004;23(20):3105–24.
14. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. Med Decis Making. 2013;33(5):607–17.
15. Siddiqui MK, Khurana IS, Budhia S, Hettle R, Harty G, Wong SL. Systematic literature review and network meta-analysis of cladribine tablets versus alternative disease-modifying treatments for relapsing-remitting multiple sclerosis. Curr Med Res Opin. 2018;34(8):1361–71.
16. McCool R, Wilson K, Arber M, Fleetwood K, Toupin S, Thom H, et al. Systematic review and network meta-analysis comparing ocrelizumab with other treatments for relapsing multiple sclerosis. Mult Scler Relat Disord. 2019;29:55–61.
17. Liu Z, Liao Q, Wen H, Zhang Y. Disease modifying therapies in relapsing-remitting multiple sclerosis: a systematic review and network meta-analysis. Autoimmun Rev. 2021;20(6):102826.
18. Drudge C, Zhao M, Tanaka S, Tanaka N, Otaka H, Kawachi I, et al. Systematic review and network meta-analysis comparing ofatumumab with other disease-modifying therapies available in Japan for the treatment of patients with relapsing multiple sclerosis. Clin Exp Neuroimmunol. 2022;13(4):234–45.
19. Upreti VV, Venkatakrishnan K. Model-based meta-analysis: optimizing research, development, and utilization of therapeutics using the totality of evidence. Clin Pharmacol Ther. 2019;106(5):981–92.
20. Chan P, Peskov K, Song X. Applications of model-based meta-analysis in drug development. Pharm Res. 2022;39(8):1761–77.
21. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097.
22. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol. 2016;75:40–6.
23. Higgins J GS. Cochrane handbook for systematic reviews of interventions version 6.3. Hoboken: Wiley Blackwell; 2022.
24. Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on indirect treatment comparisons good research practices: part 1. Value Health. 2011;14(4):417–28.
25. Cooper NJ, Sutton AJ, Morris D, Ades AE, Welton NJ. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. Stat Med. 2009;28(14):1861–81.
26. Jansen JP, Naci H. Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers. BMC Med. 2013;11(1):159.
27. Cope S, Zhang J, Saletan S, Smiechowski B, Jansen JP, Schmid P. A process for assessing the feasibility of a network meta-analysis: a case study of everolimus in combination with hormonal therapy versus chemotherapy for advanced breast cancer. BMC Med. 2014;12:93.
28. Salanti G, Marinho V, Higgins JP. A case study of multiple-treatments meta-analysis demonstrates that covariates should be considered. J Clin Epidemiol. 2009;62(8):857–64.
29. Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence synthesis for decision making 3: heterogeneity–subgroups, meta-regression, bias, and bias-adjustment. Med Decis Making. 2013;33(5):618–40.
30. Vollmer T, Signorovitch J, Huynh L, Galebach P, Kelley C, DiBernardo A, et al. The natural history of brain volume loss among patients with multiple sclerosis: a systematic literature review and meta-analysis. J Neurol Sci. 2015;357(1–2):8–18.
31. Mandema JW, Gibbs M, Boyd RA, Wada DR, Pfister M. Model-based meta-analysis for comparative efficacy and safety: application in drug development and beyond. Clin Pharmacol Ther. 2011;90(6):766–9.
32. Mandema JW, Salinger DH, Baumgartner SW, Gibbs MA. A dose-response meta-analysis for quantifying relative efficacy of biologics in rheumatoid arthritis. Clin Pharmacol Ther. 2011;90(6):828–35.
33. Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. London: National Institute for Health and Care Excellence (NICE); 2014. https://pubmed.ncbi.nlm.nih.gov/27466657.
34. Olsson T, Boster A, Fernández Ó, Freedman MS, Pozzilli C, Bach D, et al. Oral ponesimod in relapsing–remitting multiple sclerosis: a randomised phase II trial. J Neurol Neurosurg Psychiatry. 2014;85(11):1198.
35. Kappos L, Li D, Calabresi PA, O'Connor P, Bar-Or A, Barkhof F, et al. Ocrelizumab in relapsing-remitting multiple sclerosis: a phase 2, randomised, placebo-controlled, multicentre trial. The Lancet. 2011;378(9805):1779–87.
36. Giovannoni G, Comi G, Cook S, Rammohan K, Rieckmann P, Sørensen PS, et al. A placebo-controlled trial of oral cladribine for relapsing multiple sclerosis. N Engl J Med. 2010;362(5):416–26.
37. Rae-Grant A, Day GS, Marrie RA, Rabinstein A, Cree BAC, Gronseth GS, et al. Practice guideline recommendations summary: Disease-modifying therapies for adults with multiple sclerosis: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. Neurology. 2018;90(17):777–88. https://doi.org/10.1212/WNL.0000000000005347.
38. Barkhof F. The clinico-radiological paradox in multiple sclerosis revisited. Curr Opin Neurol. 2002;15(3):239–45.
39. Li DK, Held U, Petkau J, Daumer M, Barkhof F, Fazekas F, et al. MRI T2 lesion burden in multiple sclerosis: a plateauing relationship with clinical disability. Neurology. 2006;66(9):1384–9.
40. Tedeschi G, Lavorgna L, Russo P, Prinster A, Dinacci D, Savettieri G, et al. Brain atrophy and lesion load in a large population of patients with multiple sclerosis. Neurology. 2005;65(2):280–5.
41. Samjoo IA, Worthington E, Drudge C, Zhao M, Cameron C, Häring DA, et al. Comparison of ofatumumab and other disease-modifying therapies for relapsing multiple sclerosis: a network meta-analysis. J Comparative Effectiveness Research. 2020;9(18):1255–74.
42. Samjoo IA, Drudge C, Walsh S, Tiwari S, Brennan R, Boer I, et al. Comparative efficacy of therapies for relapsing multiple sclerosis: a systematic review and network meta-analysis. J Comp Eff Res. 2023;12(7):e230016.
43. Fisher E, Barkhof F, Elskamp IJ, Miller M, Jasperse B, Allen R, et al. Comparison of brain atrophy measurement methods in the context of a clinical trial. Dusseldorf: European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS); 2009.
44. Bermel RA, Bakshi R. The measurement and clinical relevance of brain atrophy in multiple sclerosis. Lancet Neurol. 2006;5(2):158–70.

45. Guevara C, Garrido C, Martinez M, Farias GA, Orellana P, Soruco W, et al. Prospective assessment of no evidence of disease activity-4 status in early disease stages of multiple sclerosis in routine clinical practice. Front Neurol. 2019;10:788.
46. Narayanan S, Nakamura K, Fonov VS, Maranzano J, Caramanos Z, Giacomini PS, et al. Brain volume loss in individuals over time: source of variance and limits of detectability. Neuroimage. 2020;214:116737.

Publisher's Note