MSJ Experimental Translational Clinical

Review

Treatment effect modifiers of immunotherapies for relapsing-remitting multiple sclerosis—A systematic review and meta-analysis

Multiple Sclerosis Journal—Experimental, Translational and Clinical

October–December 2024, 1–11

DOI: 10.1177/20552173241274618

Christoph Heesen*, Christian Röver*, Susanna Salem, Judith Heinz, Declan Chard, Jordi Rio, Andrea V Fittipaldo, Thomas Lehnert, Sascha Köpke, Alessandra Solari, Maria Pia Sormani, Tim Friede and Anne C Rahn

Abstract

Background: This meta-analysis aimed to assess the treatment effects of immunotherapies in subgroups of adults with clinically isolated syndrome or relapsing forms of multiple sclerosis (MS) and the effect of potential treatment effect modifiers (TEMs).

Methods: Phase 2 and 3 RCTs with a placebo comparator were analyzed. Risk of bias was assessed. Random-effects meta-analyses were conducted to summarize treatment effects within subgroups and differences in treatment effects between subgroups.

Results: Thirty-one studies were included. Age < 40 years was the strongest TEM for relapse rate across DMTs with a ratio of rate ratios (RRR) of 1.44 (95% CI 1.09–1.90; 7 studies). Disability progression was influenced by age (ratio of hazard ratios, RHR 1.59, 95% CI 1.11–2.29; 4 studies). Dichotomizing patients based on EDSS cut-offs (EDSS 2.0 and 3.0) also showed a significantly higher benefit for those less disabled for relapse rate (RRR 1.35, CI 1.03–1.76; 8 studies). Sex, baseline MRI parameters, previous immunotherapy, and clinical presentation showed no effect in this meta-analysis.

Conclusion: Age < 40 is a robust TEM for a lower relapse rate as well as less disability progression across six MS immunotherapies. Additionally, a lower baseline EDSS was predictive of the relapse rate.

Keywords: Multiple sclerosis, disease-modifying therapy, treatment effect modifier, treatment response, meta-analysis, systematic review

Date received: 7 March 2024; accepted 25 July 2024

Introduction

Over the last three decades, more than 15 disease-modifying therapies (DMTs) with different mechanisms of action and benefit–risk profiles have been approved for treating relapsing-remitting multiple sclerosis (RRMS).[1,2] They reduce relapse rates and possibly slow disability accrual and overall disease progression.[2–6] Evidence from randomized-controlled trials (RCTs) and observational studies suggests that starting a DMT soon after diagnosis potentially leads to better long-term outcomes, such as a reduced risk of relapses and transition to secondary progressive MS (SPMS), compared to later initiation.[7] This favorable effect seems more pronounced for DMTs with higher efficacy.[8,9] But DMTs with higher efficacy also exhibit higher toxicity, which in rare cases may lead to severe and potentially life-threatening adverse effects.[10,11]

While an optimal treatment window for starting and escalating immunotherapies may close in the disease course when degenerative processes dominate, the reliable identification of patients with poor responses is crucial.[12]

Treatment effects are not homogeneous across trial participants, as suggested by findings from subgroup analyses of individual RCTs[13,14] but vary according

*Equally contributing authors.

Correspondence to:
Christoph Heesen, Institute of Neuroimmunology and Multiple Sclerosis (INIMS), University Medical Center, Martinistr. 52, 20246 Hamburg, Germany.
heesen@uke.de

Christoph Heesen, Institute of Neuroimmunology and Multiple Sclerosis (INIMS), University Medical Center, Hamburg, Germany

Department of Neurology, University Medical Center, Hamburg, Germany

Christian Röver, Susanna Salem, Judith Heinz, Department of Medical Statistics, University Medical Center Göttingen, Göttingen, Germany

Declan Chard, Department of Neuroinflammation, UCL Queen Square Institute of Neurology, Faculty of Brain Sciences, University College London, London, UK National Institute for Health Research (NIHR), University College London Hospitals (UCLH) Biomedical Research Centre, London, UK

Jordi Rio, Neurology/ Neuroimmunology, Centre d'Esclerosi Multiple de Catalunya (Cemcat), Hospital Universitari Vall d'Hebron, Universitat Autonoma de Barcelona, Barcelona, Spain

Andrea V Fittipaldo, Department of Oncology, Istituto Ricerche Farmacologiche "Mario Negri" IRCCS, Milano, Italy

Thomas Lehnert, Institute of Neuroimmunology and Multiple Sclerosis (INIMS), University Medical Center, Hamburg, Germany

Sascha Köpke, Institute of Nursing Science, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

Alessandra Solari, Unit of Neuroepidemiology, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy

Maria Pia Sormani, Department of Health Sciences, University of Genua, Genova, Italy

Tim Friede, Department of Medical Statistics, University Medical Center Göttingen, Göttingen, Germany

Anne C Rahn, Institute for Social Medicine and Epidemiology, Nursing Research Unit, University of Lübeck, Lübeck, Germany

to demographic and disease-related characteristics.[15,16] One method to identify poor responders lies in the assessment of treatment effect modifiers (TEMs) in RCTs. Prognostic and treatment response predictions are controversial in MS.[14,17] Evidence, however, indicates that higher treatment benefit is observed in younger persons with MS (pwMS), and with lower disability, and less inflammatory disease activity as shown on MRI.[13,16]

While Signori et al.[13] performed a meta-analysis on all published phase-3 trials with available TEM analysis until 2014, to the best of our knowledge, there has been no updated systematic review considering TEMs in all the presently available DMT options for clinically isolated syndrome (CIS) or RRMS in comparison to placebo.

Methods

This systematic review aimed to assess differential treatment effects (responses) to approved immunotherapies in subgroups of adults with CIS or RRMS. We registered the review[18] and published the protocol.[19]

Inclusion criteria

We included all published primary and secondary analyses of phase-2 and phase-3 RCTs, including randomized crossover trials. Participants were adult patients with CIS or RRMS diagnosis based on Poser or McDonald criteria. Follow-up had to be at least 12 months. Studies had to compare at least one approved DMT to a placebo. Subgroup characteristics needed to have been assessed at baseline, i.e. before participants were exposed to the study drug. We considered subgroups defined by age, disability level, sex, relapse history, MRI activity and treatment.

Search strategy

The following sources were searched (last accessed 31 July 2022): Cochrane Central Register of Controlled Trials (CENTRAL 2020), MEDLINE (PubMed, 1946 to search date), Embase (Embase.com, 1974 to search date), Cumulative Index to Nursing and Allied Health Literature (CINAHL, EBSCO host, 1981 to search date), Latin American and Caribbean Health Science Information Database (LILACS, 1982 to search date), ClinicalTrials.gov (http://clinicaltrials.gov), World Health Organization (WHO) International Clinical Trials Registry Platform (https://trialsearch.who.int/). The search keywords are listed in Appendix 1.

Screening and data extraction

We used Covidence (https://www.covidence.org) for the screening process. Two review authors (AF, TL, CH, AR) performed the screening independently. Disagreements were resolved by discussion. All potentially relevant reports were retrieved in full-text and investigated for inclusion independently by two authors (CH, AR). Disagreements were again solved through discussion. Reasons for study exclusion were recorded as given in the PRISMA[20] flow chart (see Figure 1).

Data extraction of included studies was performed by one author (SS) based on a piloted excel form, and checked by a second author (JH).

Assessment of risk of bias

We used the Cochrane "Risk of Bias 2" tool (RoB2[21]) for RCTs to assess the included studies, inclusive of the main publications concerning the secondary analyses. Assessments were conducted independently by two authors (CH, AR) for each outcome. Disagreements were resolved by discussion. As the results differed according to different outcomes we analyzed risk of bias data for every outcome and aggregated the findings in Appendix Table 4.

Measures of treatment effect

Primary outcomes included relapses, disability progression, quality of life measures and adverse events. Secondary outcomes included MRI parameters, evolution of no-evidence-of-disease-activity (NEDA) 3 criteria, the activity of daily living, fatigue, conversion to clinically definite MS and SPMS.[22]

Treatment effects on binary endpoints (e.g. confirmed disability worsening) were analyzed as odds ratios (ORs), based on the proportions of affected patients. Treatment effects on rates (e.g. relapses, adverse events) were analyzed using rate ratios (RRs) when event rates were reported or computable, or using ORs when only proportions of affected patients were reported. Treatment effects expressed in terms of continuous endpoints (e.g. time to CDMS) were analyzed based on mean differences. Time-to-event endpoints were analyzed using hazard ratios (HRs). Estimates of overall mean effects were quoted along with 95% confidence intervals (CIs). In case results were reported both for ITT and per-protocol analysis sets, the ITT results were preferred. The main focus of this study is on TEMs, i.e. the differing treatment response between patient subgroups. These interaction effects are expressed in terms of ratios of rate

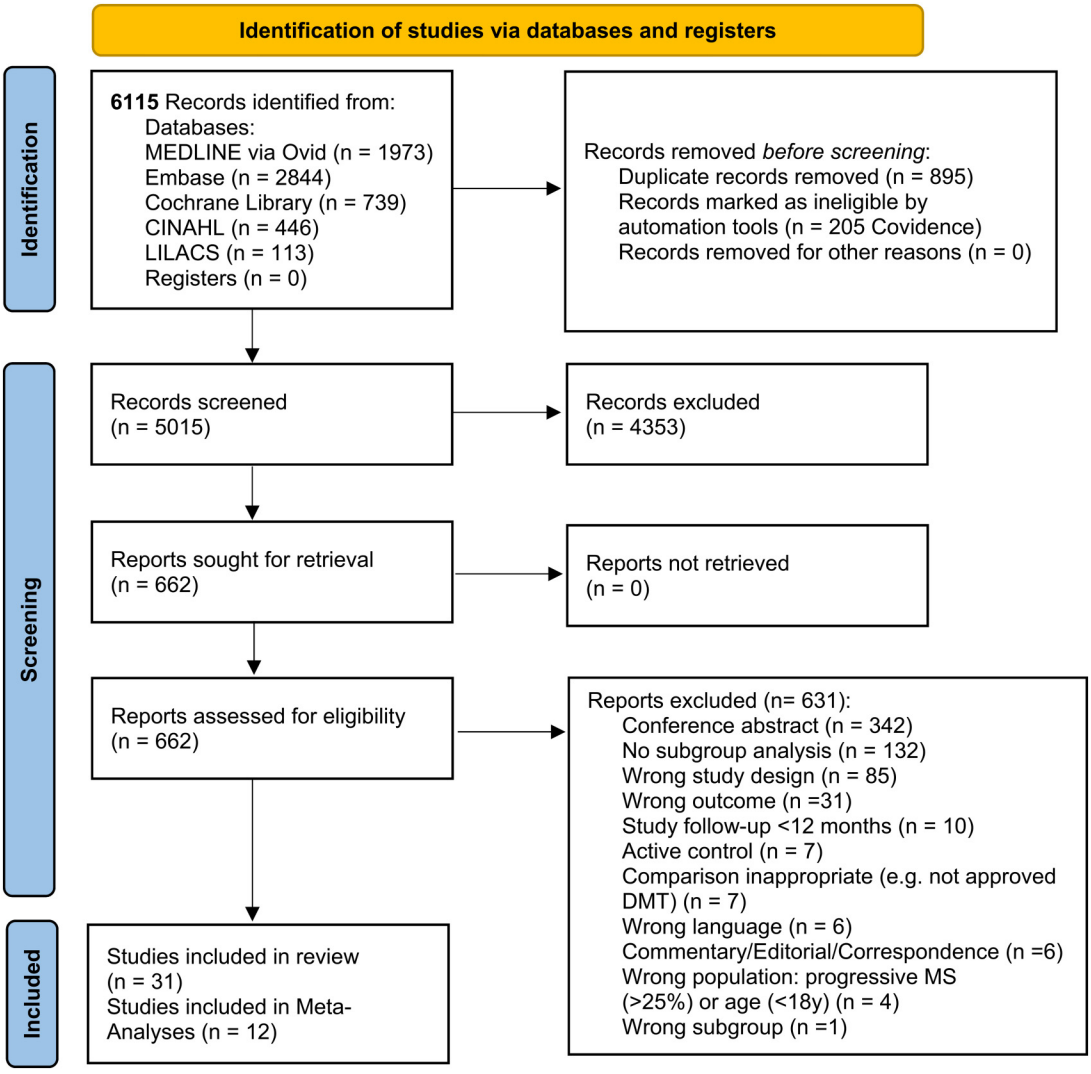


Figure 1. PRISMA flow-chart of study selection. The 31 studies (for overview and references see Table S1) included the following DMTs: 14 interferon-beta preparations, 2 glatiramer acetate, 4 dimethlyfumarate, 3 Natalizumab, 2 sphingosine-1-phosphate receptor modulator, 2 Teriflunomide, 4 Cladribine. In total, 9 studies addressed CIS and 24 studies addressed RRMS. In total, 11 reported relapses, and 10 studies reported EDSS or relapses as an outcome. Nine studies reported MRI findings and four studies reported NEDA as an outcome. An overview of all studies see Appendices 2 and 3.

ratios (RRRs), ratios of odds ratios (RORs), etc. (see also Figure 2 for an illustration). Results are displayed as forest plots.

Assessment of heterogeneity

Between-trial heterogeneity was accounted for using random-effects models, where the heterogeneity parameter (the standard deviation tau) is estimated using the Paule-Mandel method.[23] We estimated heterogeneity (tau) on its absolute scale. Heterogeneity estimates are indicated at the bottom of forest plots. In case there was evidence for the presence of "fairly extreme" heterogeneity,[24] considering the heterogeneity estimate and its 95% CI, we refrained from a quantitative synthesis and reported a qualitative summary only.

Data synthesis

If studies showed sufficient similarity in participants, eligibility criteria and outcomes, we pooled effect estimates from all DMTs for the overall trial population with placebo as a control and conducted meta-analyses for each relevant outcome. Following the procedure for the overall trial population, we pooled effect estimates from all DMTs, but this time for each subgroup with placebo as a control and

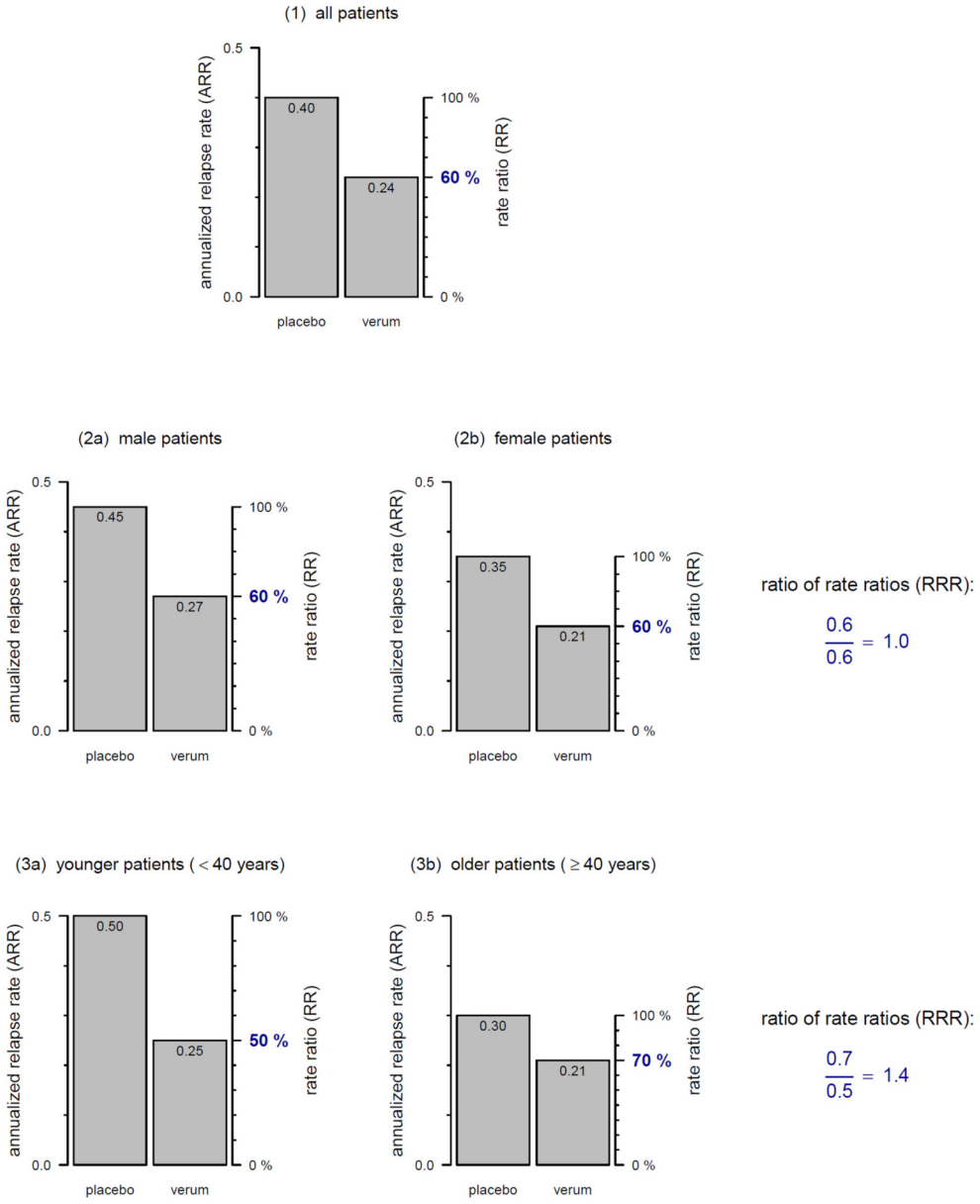


Figure 2. Calculation of interaction effects. The figure illustrates the presence or absence of a treatment-by-subgroup interaction based on a fictitious clinical trial. Suppose a treatment (verum) was investigated in a placebo-controlled trial, the outcome of interest being the (relative) reduction in the annualized relapse rate (ARR). Besides the overall study population, analyses also focused on sex (male/female) and age (younger/older) subgroups of patients. Overall, the ARR in the placebo group is 0.4 relapses per patient per year; in the treatment group the rate is reduced to only 60% of the original rate (i.e. a reduction by 40%), corresponding to a rate ratio (RR) of 0.6 (see panel 1). The *y*-axis on the left-hand side of the plot indicates the annualized relapse rates (ARR) for both treatment groups. The *y*-axis on the right-hand side shows the reduction in the event rate (the rate ratio) relative to the control group as the reference. When considering male and female patients (panels 2a and 2b), there is a slight difference in the ARRs under placebo: men tend to experience more relapses than women. When considering the treatment effect, however, the relative effect still remains constant: in both subgroups, the rate is reduced to 60% compared to placebo. With differing placebo rates, an identical relative effect here means that the absolute reduction in event rate also differs between patient subgroups. With rate ratios (RRs) of 60% in both groups, the ratio of RRs (RRR) now results as 0.6/0.6 = 1.0, indicating no treatment-by-subgroup interaction. When analyzing age, a similar picture emerges (panels 3a and 3b). The relapse rates differ between subgroups, with younger patients having more relapses. However, the relative treatment effects now also differ between subgroups: for young patients, the treatment reduces the rate to 50%, while for older pwMS a reduction to 70% is achieved. For the age subgroups, the RRR hence is at 0.7/0.5 = 1.4, indicating a treatment-by-subgroup interaction.

conducted meta-analyses. This analysis preceded the main analyses and illustrates the (combined) interventions' overall effects. Analyses were carried out using random-effects models based on a normal-normal hierarchical model, combining effect (or interaction) estimates using generic inverse-variance methods and Paule-Mandel estimates for the heterogeneity variance. We used "R" and the "metafor" and "forestploter" packages.[25,26] In addition to the "classical" (Wald type) confidence intervals, we also provide Hartung–Knapp–Sidik–Jonkman (HKSJ) intervals for the interaction estimates in Appendices 6a and 6b. When meta-analysis was not possible, single studies were not presented.

For the TEMs baseline Expanded Disability Status Scale Scores (EDSS[27] (cut-offs ≤2.0 or >2.0 and ≤3.5 or >3.5) as well as for the number of relapses in the 12 to 24 months before study start different cut-offs were pooled in an additional analysis to gain more information. MRI baseline lesion stratification was based on the four dissemination in space criteria in the 2005 revision of McDonald.[28] These are: (a) at least one gadolinium-enhancing lesion or nine T2 hyperintense lesions if there is no gadolinium-enhancing lesion, (b) at least one infratentorial lesion, (c) at least one juxtacortical lesion, and (d) at least three periventricular lesions.

Results

The systematic literature search resulted in 6115 references to be screened, of which 662 full-texts were analyzed. Thirty-one studies were included. Figure 1 displays screening and exclusion steps and Appendix 2 gives an aggregated overview of included studies. Appendices 3a and 3b represent single studies including interactions and applied cut-offs. Twelve studies contributed to the final meta-analyses (Figure 1). Due to different endpoints and different estimates (either hazard, rate or odds ratios), it was not always possible to combine different studies on a given measure of treatment effect modification.

Risk of bias

Overall, the methodological quality of the studies was high (see Appendix 4). However, for 25 studies, the analysis plan was not published in advance. Only ten studies had an additional risk of bias, the most common being lack of reliable blinding of subjects or raters, and no evidence of rater training. None of these was severe enough to exclude a study from the analysis.

Presentation of the effects of TEM in detail

We provide forest plots on all results in Appendices 5a (CIs Wald-type) and 5b (CIs HKSJ-type). RHR refers to the ratio of hazard-ratios, ROR refers to the ratio of odds ratios, RRR to the ratio of rate ratios.

For the primary outcomes quality of life and adverse events and the secondary outcomes evolution of NEDA3, activity of daily living, fatigue, conversion to SPMS only single or no studies were available, therefore no meta-analyses were performed.

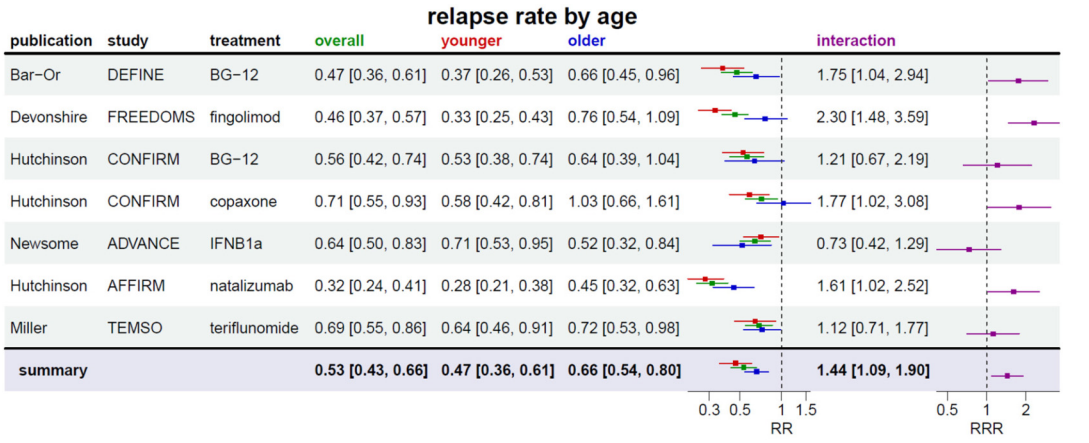
Age. Age below 40 and above or equal to 40 years was examined in the TEM analysis as 40 years was the most commonly applied cut-off value in the source studies (see Figure 3). Here, we found the strongest effects in the entire analysis. For the endpoint relapses, we found an RRR of 1.44 (95% CI 1.09–1.90, seven trials) and RHR of 1.68 (95% CI 1.23–2.28, three studies) indicating a smaller treatment effect in older persons with MS (pwMS). Analyses on the endpoint of disability progression showed an RHR of 1.63 (CI 1.15–2.31, four studies). The interaction effect was consistent in most studies, except for the ADVANCE study [54],[9] suggesting an inverse effect regarding the age groups.

Age did not show differences in treatment effects on the endpoints: T2 lesion count (RRR 0.92, 95% CI 0.55–1.56, two studies), number of gadolinium-enhancing lesions (ROR 0.82, 95% CI 0.3–2.28, two studies) or time to CDMS (RHR 1.24, 95% CI 0,8–1.93, three studies).

Sex. We found no significant interaction effects regarding sex as TEM. For sex, the effect of the treatment on the endpoint relapse rate had a non-significant RRR of 1.10 (95% CI 0.90–1.36, seven studies) and a non-significant RHR for time to relapse (RHR 1.22, 95% CI 0.87–1.7, three studies).

The treatment effects on the endpoint time to disability progression between males and females did as well not differ, reflected in a RHR of 1.30 (95% CI 0.88–1.93, four studies). We also found no effects for T2 lesion number (RRR 0.96, 95% CI 0.56–1.62, two studies), gadolinium enhancement (ROR 0.45, 95% CI 0.17–1.17, two studies) or time to development of CDMS (RHR 0.92, 95% CI 0.58–1.48, three studies).

Relapse rate. The treatment effect on the relapse rate ratio during the study was not associated with the relapse rate in the year before the study (RRR 0.91,



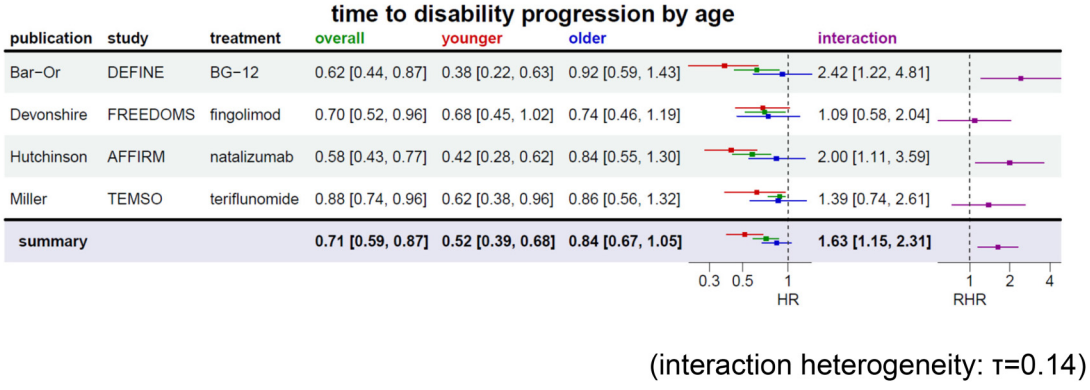


Figure 3. Age as TEM for the outcomes relapse rate and disability progression. Both meta-analyses showed significant effects for age below or above 40 as a TEM.

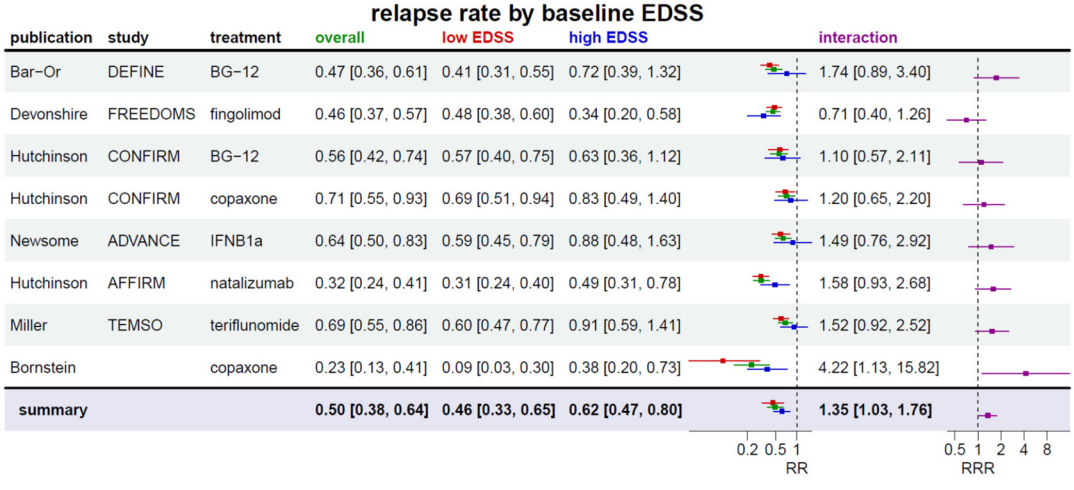


Figure 4. Baseline EDSS as a TEM for on-study relapse rate. Cut-offs of >3.5 and >2.0 were aggregated and compared to cut-offs ≤3.5 and ≤2.0.

95% CI 0.72–1.15, five studies) or within two preceding years before study entry (RHR 0.58, 95% CI 0.31–1.08, two studies). There was no association of pre-study relapses in the year before starting treatment with the effect on disability progression (RHR 0.92, 95% CI 0.62–1.36, three studies). The same applies for relapse rate within 2 years before study start (RHR 0.58, 95% CI 1.31–1.08, two studies). Also, there were no treatment effects on T2 lesion number (RRR 1.49, 95% CI 0.59–3.78, two studies) and gadolinium accumulation (ROR 1.43, 95% CI 0.53–3.90, two studies) modified by prestudy relapse rates. Aggregating the subgroup data regarding previous relapses over 1 and 2 years before the studies began did not change the estimates for the treatment effects on relapse rate or disability (RRR 0.93, 95% CI 0.76–1.14, two studies).

Baseline EDSS. Disability at study inclusion, grouped according to a cut-off value of 3.5 showed a RRR of 1.29 (95% CI 1.03–1.62, seven studies) for the treatment effect on relapses (see Figure 4). For the cut-off 2.0, the RRR was 1.76 (95% CI 1.12–2.77, four studies) for the treatment effect on relapses.

The RHR for the treatment effect on disability progression was not significant for the 3.5 cut-off (RHR 0.81, 95% CI 0.39–1.66, four studies) as well as for the 2.0 cut-off (RHR 1.41, 95% CI 0.98–2.04, three studies). Aggregating the EDSS cut-offs, however, showed a significant effect on the treatment effect on relapse rate (RRR 1.35, 95% CI 1.03–1.76, eight studies, Figure 4) suggesting a stronger treatment benefit for less severely disabled patients.

We found no significant differences for the endpoint disability for the cut-off 3.5 (RHR 0.81, 95% CI 0.39–1.66, four studies). The treatment effect on gadolinium-enhancing lesions (ROR 0.93, 95% CI 0.38–2.29, two studies) or the T2 lesion load (RRR 1.13, 95% CI 0.68–1.87, two studies) was not associated with the baseline EDSS cut-off 2.0.

T2 lesion burden (MRI, volume or number). The T2 lesion burden, grouped by above or below median lesion volume, was not associated with the treatment effect on relapse rate (RRR 1.05, 95% CI 0.80–1.37, five studies). As well time to relapse was not influenced by T2 lesion volume (RHR 0.84, 95% CI 0.54–1.31, three studies). There was also no association of T2 volume on the treatment effect on disability (RRR 0.64, 95 CI 0.38–1.09, two studies). No effects were found concerning the evolution of gadolinium accumulation (ROR 0.55, 95% CI 0.22–1.40, two studies) or T2 lesion volume evolution (RRR 0.84, 95% CI 0.50–1.40, two studies) when differentiating T2 volume above or below median at baseline. The treatment effect on the conversion to CDMS was not influenced by having less or equal to or more than 9 T2 lesions (RHR 1.25, 95% CI 0.76–2.05, four studies).

Gadolinium enhancement (MRI). The presence of at least one Gd-positive lesion did not show any association with the treatment effect on relapse rate (RRR 0.86, 95% CI 0.70–1.05, seven studies). Similarly, time to relapse was not influenced by Gd-positivity (RHR 0.97, 95% CI 0.63–1.52, three studies). Furthermore, the presence of Gd-positive lesions was also not found to be associated with the treatment effect on time to disability progression (RHR 0.75, 95% CI 0.53–1.06, four studies). The treatment effect on the development of CDMS was not dependent on the presence of at least one gadolinium-enhancing lesion at baseline (RHR 0.74, 95% CI 0.49–1.12, five studies). There was also no effect of baseline Gd-enhancing lesions on the subsequent number of T2 lesions on treatment (RRR 0.67, 95% CI 0.27–1.64, two studies). And a positive Gd-status at baseline did not influence the appearance of Gd-positive lesions at follow-up (ROR 0.35, 95% CI 0.18–0.86, two studies).

McDonald criteria. The comparison of McDonald-1 (one Gd + lesion or nine periventricular lesions) versus McDonald 2–4 (one or more infratentorial or juxtacortical lesion or more than three periventricular lesions) showed no interaction with treatment effect on the endpoint relapse rate (RRR 0.91, 95% CI 0.61–1.35, four studies) or time to relapse (RHR 0.87, 95% CI 0.57–1.34, three studies). Different McDonald criteria (1 or 2,3, 4) did not impact the T2 lesion number (RRR 1.04, 95% CI 0.48–2.28, two studies) or Gd-status on treatment (ROR 1.17, 95% CI 0.38–3.65, two studies).

Previous immunotherapy. Treatment with one DMT or more than one had no impact on the treatment effect on relapse rate (RRR 1.10, 95% CI 0.82–1.47, three studies), time to relapse (RHR 1.06, 95% CI 0.64–1.74, three studies) or time to disability progression (RHR 1.14, 95% CI 0.55–1.36, four studies). T2 lesion number (RRR 0.82, 95% CI 0.49–1.36, two studies) as well as number of Gadolinium-enhancing lesions (ROR 1.22, 95% CI 0.2–7.42, two studies) were as well not influenced by previous treatment.

Monofocal or multifocal presentation. For the type of clinical presentation, comparable studies were found for the endpoint of clinically definite MS in studies treating clinically isolated syndromes. Here, no statistically robust treatment effect modification for a monofocal of multifocal presentation was found (RHR 1.23, 95% CI 0.79–1.92, three studies).

T1 lesions. No comparable studies were found for T1 lesions. Therefore, a meta-analysis could not be could be carried out.

In Appendix 5b, we provide additional HKSJ intervals for the interaction estimates of all analyses.

Discussion

This is, to the best of our knowledge, the largest systematic review and meta-analysis of TEMs in MS. Overall, it reaffirms previous findings that age <40 is a robust TEM for a greater reduction in relapse rate as well as less disability progression across six different MS immunotherapies.

Overall age influenced treatment response in terms of both relapse rate and time-to-disability-progression, with a substantially greater benefit for younger patients (up to 50% relative reduction in pwMS ≤ 40 years). This is consistent with a previous meta-analysis. Signori et al. studied subgroups from six RCTs and demonstrated age, followed by baseline disability and gadolinium-enhancing lesions as TEMs.[13] In our analysis, the contributing fingolimod study seemed to have the strongest interaction with age (Appendix 5.1). While this comparison across studies is limited, this age dependency is also well reflected in the high efficacy data of fingolimod in children showing an absolute risk reduction by 46% percentage points for being relapse-free compared to interferon-beta.[29] Also, in accordance with other data, the age-dependency of teriflunomide effects seems rather weak, which is mirrored in the same effect of teriflunomide in children as in adults.[30] Weidemann et al.[16] came to similar conclusions in their meta-analysis on age. Taken together, this suggests that immunotherapy is less likely to be of benefit with higher age and should be considered carefully in the context of risk profile.

Second to age, baseline disability, based on an EDSS cut-off 2.0 or 3.5, showed an impact on the treatment effect on relapse rate. Age effects might partially explain this, as with increasing age EDSS increases as well. However, studying the impact of baseline EDSS on disability evolution only four studies could be aggregated. Aggregating data from dimethyl fumarate, natalizumab or teriflunomide studies could not show a modification of treatment effect on disability by a baseline EDSS cut-off of 3.5. This is in line with the meta-analysis of Signori et al., reporting a non-significant impact of baseline EDSS on disability evolution.[13]

Our results also indicate that neither a high lesion burden, the detection of Gd-enhancing lesions nor a high relapse rate before the start of therapy predict the response to immunotherapy. Consistent with our findings, Pellegrini et al. found age (and quality of life) as the strongest TEM but no MRI parameter in a recent modeling study based on CONFIRM and DEFINE.[31]

Sex did not predict treatment response in our meta-analysis, although descriptively, the male sex was rather associated with more disability evolution (Appendix 5). While Signori et al.[13] did not report on sex effects, Li et al.[32] also did a systematic meta-analysis showing no clear association between sex and treatment response in 16 studies.

As no raw data were available we could only perform meta-analysis of datasets of at most seven studies for a TEM, and in most cases only three to four studies could be aggregated. Subgroup analysis is usually not part of a core study design and analysis concept leading to a lack of power to detect effects. In addition, subgroup sizes are usually not balanced within a study, unless a cut-off is pragmatically drawn at or near the median. This also explains why findings can differ even after looking at the same DMT in different studies. However, the pattern seems relatively consistent within the substance groups of interferon-beta preparations. Confounders of a TEM, such as disease duration or age, for disability have not been investigated in the TEM studies. Thus certain confounders may constitute the actual causes of the observed patterns. Furthermore, even trial durations of up to three years are short in the context of progression seen in natural history studies, and longer-term follow-ups may be necessary to enable a definitive appraisal of a TEM. It is worth also noting that because of the large amount of available data we did not include the studies comparing different immunotherapies.

Random-effects models were used for all meta-analyses, which may become unreliable when few studies are included. In the present project, the majority of analyses were based on two to three studies, and other analysis methods, such as HKSJ,

have been proposed that may reduce the risk of type I errors, albeit at the expense of wider confidence intervals leading to inconclusive results.[33] In the present study, the use of the HKSJ approach would have resulted in the apparent significance of 5 of the 51 results changing (4 becoming non-significant, while 1 became significant; see Appendix 5b for the adjusted results).[33]

Finally our meta-analysis integrated findings from studies of immunotherapies with very different modes of action, and while they all reduce the accrual of white matter lesions and risk of relapses, this may lead to differences in the potential for TEM to impact on outcomes. With the present data, it is not possible to robustly look for differences in TEMs between drugs, given other methodological differences (e.g. cohort recruitment criteria and how treatment outcome measures were assessed) in the studies that may confound comparisons.

In contrast to our systematic analysis of phase-2 and phase-3 RCTs, larger scale registry data found TEM potential of all the factors we have studied here.[34] While data from registries and RCTs may complement each other to some degree, both commonly differ vastly in size and data quality, and in particular, randomization and blinding enhance the reliable estimation of causal treatment effects in an RCT. Kalincik et al.[34] confirmed their findings in an independent cohort from the Swedish MS registry, but we are unaware of other registry-based treatment prediction studies.

In conclusion, this meta-analysis confirms that age is a significant TEM in the treatment with interferon-beta, glatiramer actetat, dimethyl fumrate, teriflunomide and natalizumab. Inflammatory activity (lesion load, Gd-enhancing lesion counts and relapses) at baseline or within one or two years before treatment initiation were not significant predictors of therapeutic responses.

Review protocol

Lehnert T, Röver C, Köpke S, et al. Immunotherapy for people with clinically isolated syndrome or relapsing-remitting multiple sclerosis: treatment response by demographic, clinical, and biomarker subgroups (PROMISE)—a systematic review protocol. *Syst Rev* 2022;11(1):134. doi: 10.1186/s13643-022-01997-2. PMID: 35778721.

Acknowledgements

The authors would like to acknowledge the contributions of Edeltraud Faßhauer, Claudia Schilewski and Markus van der Loo. Their work as consumer representatives helped to define outcomes and progress the protocol.

Author contributions

CH, CR, DC, MPS, AR drafted and edited the paper. All authors reviewed and approved the protocol before publication.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Federal Ministry of Education and Research (grant number 01KG1804).

Declaration of conflicting interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article:

JR: has received honoraria for consultancy from Mylan, Novartis and Sanofi-Aventis and compensation for lectures and educational presentations from, Merck- Serono, Novartis, Teva, and Sanofi-Aventis

DC: is a consultant for Biogen and Hoffmann-La Roche. In the last three years he has received research funding from Hoffmann-La Roche, the International Progressive MS Alliance, the MS Society, and the National Institute for Health Research (NIHR) University College London Hospitals (UCLH) Biomedical Research Centre, and speaker's honorarium from Novartis. He co-supervises a clinical fellowship at the National Hospital for Neurology and Neurosurgery, London, which is supported by Merck.

MPS: has received Grants (Roche, Novartis, Biogen), Fees for educational events and lectures (Roche, Novartis, Biogen, Merck, Sanofi, Celgene), Support for travelling and participation to meetings (Roche, Novartis, Biogen, Merck, Sanofi, Celgene), Payment for participation on a Data Safety Monitoring Board, Advisory Board, or Guideline Panel (Roche, Novartis, Biogen, Merck, Sanofi, Celgene, GeNeuro, Medday), and that they have been involved in conducting a study funded by Merck that is eligible for inclusion in the review.

TF: has received within the past three years personal fees for participation in the Scientific Board and lectures of the JUMPstart programme from Fresenius Kabi, for participation in data monitoring committees from Novartis, Bayer, Janssen, Roche, Daiichi-Sankyo, Boehringer Ingelheim, Coherex Medical and BiosenseWebster, and for statistical consultancies from Novartis, Bayer, CSL Behring, Galapagos, Vifor, Mediconomics, Penumbra, but all outside the submitted work.

CH: has received research grants from Roche, Merck, Bristol Myers Squibbs, Sanofi, honoraries for talks from Merck and a travel grant from Celgene. Remaining authors have none declared.

Ethics approval

Not applicable.

Review registration

Prospero 2021 CRD42021279665. Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021279665

ORCID iDs

Christoph Heesen https://orcid.org/0000-0001-8131-9467
Declan Chard https://orcid.org/0000-0003-3076-2682
Jordi Rio https://orcid.org/0000-0003-4546-7627
Alessandra Solari https://orcid.org/0000-0001-9930-7579
Maria Pia Sormani https://orcid.org/0000-0001-6892-104X
Tim Friede https://orcid.org/0000-0001-5347-7441

Supplemental material

Supplemental material for this article is available online.

References

1. Jakimovski D, Bittner S, Zivadinov R, et al. Multiple sclerosis. *Lancet* 2024; 403: 183–202.
2. 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.
3. Brown JWL, Coles A, Horakova D, et al. Association of initial disease-modifying therapy with later conversion to secondary progressive multiple sclerosis. *JAMA* 2019; 321: 175–187.
4. 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.
5. Coret F, Pérez-Miralles FC, Gascón F, et al. Onset of secondary progressive multiple sclerosis is not influenced by current relapsing multiple sclerosis therapies. *Mult Scler J – Exp Transl Clin* 2018; 4: 20552173 18783347.
6. Sormani MP and Bruzzi P. Can we measure long-term treatment effects in multiple sclerosis? *Nat Rev Neurol* 2015; 11: 176–182.
7. Chalmer TA, Baggesen LM, Nørgaard M, et al. Early versus later treatment start in multiple sclerosis: a register-based cohort study. *Eur J Neurol* 2018; 25: 1262–e110.
8. Harding K, Williams O, Willis M, et al. Clinical outcomes of escalation vs early intensive disease-modifying therapy in patients with multiple sclerosis. *JAMA Neurol* 2019; 76: 536–541.
9. He A, Merkel B, Brown JWL, et al. Timing of high-efficacy therapy for multiple sclerosis: a retrospective observational cohort study. *Lancet Neurol* 2020; 19: 307–316.
10. McGinley MP, Goldschmidt CH and Rae-Grant AD. Diagnosis and treatment of multiple sclerosis: a review. *JAMA* 2021; 325: 765–779.
11. Ayrignac X, Bilodeau P-A, Prat A, et al. Assessing the risk of multiple sclerosis disease-modifying therapies. *Expert Rev Neurother* 2019; 19: 695–706.
12. Gasperini C, Prosperini L, Tintoré M, et al. Unraveling treatment response in multiple sclerosis: a clinical and MRI challenge. *Neurology* 2019; 92: 180–192.
13. Signori A, Schiavetti I, Gallo F, et al. Subgroups of multiple sclerosis patients with larger treatment benefits: a meta-analysis of randomised trials. *Eur J Neurol* 2015; 22: 960–966.
14. Sormani MP and De Stefano N. Defining and scoring response to IFN-beta in multiple sclerosis. *Nat Rev Neurol* 2013; 9: 504–512.
15. Li R, Sun X, Shu Y, et al. Sex differences in outcomes of disease-modifying treatments for multiple sclerosis: a systematic review. *Mult Scler Relat Disord* 2017; 12: 23–28.
16. Weideman AM, Tapia-Maltos MA, Johnson K, et al. Meta-analysis of the age-dependent efficacy of multiple sclerosis treatments. *Front Neurol* 2017; 8: 577.
17. Mowry EM. Natural history of multiple sclerosis: early prognostic factors. *Neurol Clin* 2011; 29: 279–292.
18. Rahn A, Lehnert T, Röver C, et al. Immunotherapy for people with clinically isolated syndrome or relapsing-remitting multiple sclerosis—treatment response by demographic, clinical, and biomarker subgroups (PROMISE) – a systematic review protocol. *Prospero*. CRD42021279665. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021279665 (2021).
19. Lehnert T, Röver C, Köpke S, et al. Immunotherapy for people with clinically isolated syndrome or relapsing-remitting multiple sclerosis: treatment response by demographic, clinical, and biomarker subgroups (PROMISE)—a systematic review protocol. *Syst Rev* 2022; 11: 134.
20. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015; 4: 1.
21. Sterne JAC, et al. Rob 2: a revised tool for assessing risk of bias in randomised trials. *Br Med J* 2019; 366: l4898.
22. Lublin FD, Reingold SC, Cohen JA, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology* 2014; 83: 278–286.
23. Veroniki AA, Jackson D, Viechtbauer W, et al. Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Res Synth Methods* 2016; 7: 55–79.

24. Spiegelhalter DJ, Abrams KR and Myles JP. *Bayesian approaches to clinical trials and health-care evaluation. (Sec. 5.7)*. Chichester: Wiley, 2004.
25. R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2020.
26. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010; 36: 1–48.
27. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983; 33: 1444–1452.
28. Polman CH, Reingold SC, Edan G, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald criteria". *Ann Neurol* 2005; 58: 840–846.
29. Chitnis T, Arnold DL, Banwell B, et al. Trial of fingolimod versus interferon beta-1a in pediatric multiple sclerosis. *N Engl J Med* 2018; 379: 1017–1027.
30. Chitnis T, Banwell B, Kappos L, et al. Safety and efficacy of teriflunomide in paediatric multiple sclerosis (TERIKIDS): a multicentre, double-blind, phase 3, randomised, placebo-controlled trial. *Lancet Neurol* 2021; 20: 1001–1011.
31. Pellegrini F, Copetti M, Bovis F, et al. A proof-of-concept application of a novel scoring approach for personalized medicine in multiple sclerosis. *Mult Scler J* 2020; 26: 1064–1073.
32. Li R, Sun X, Shu Y, et al. Sex differences in outcomes of disease-modifying treatments for multiple sclerosis: A systematic review. *Mult Scler Relat Disord febbraio* 2017; 12: 23–28.
33. Jackson D, Law M, Rücker G, et al. The Hartung–Knapp modification for random-effects meta-analysis: a useful refinement but are there any residual concerns? *Stat Med* 2017; 36: 3923–3924.
34. Kalincik T, Manouchehrinia A, Sobisek L, et al. Towards personalized therapy for multiple sclerosis: prediction of individual treatment response. *Brain* 2017; 140: 2426–2443.