

ORIGINAL ARTICLE

Recommendations for the critical reading of clinical trials on disease-modifying drugs for multiple sclerosis



A. Rivero-de-Aguilar^{a,b}, M. Mascareñas-García^{b,c}, M. Pérez-Ríos^{b,d,e,*},
M.A. Llaneza-Gonzalez^f, P. Mulero^g, M. Mendibe^{h,i}, À. Rovira^j, V. Meca-Lallana^k,
A.J. García-Ruiz^l, L. Landete^m, J.R. Vizoso Hermidaⁿ, P. Carrascal Rueda^o,
A. Ruano-Raviña^{b,d,e}, M. Puente-Hernandez^p, L. Varela-Lema^{b,d,e}

^a Department of Neurology, University Hospital Complex of Pontevedra, Pontevedra, Spain

^b Department of Epidemiology and Public Health, University of Santiago de Compostela, Santiago de Compostela, Spain

^c Department of Preventive Medicine, University Hospital Complex of Santiago de Compostela, Santiago de Compostela, Spain

^d Health Research Institute of Santiago de Compostela (Instituto de Investigación Sanitaria de Santiago de Compostela – IDIS), Santiago de Compostela, Spain

^e Consortium for Biomedical Research in Epidemiology and Public Health (CIBER en Epidemiología y Salud Pública/CIBERESP), Madrid, Spain

^f Department of Neurology, Central University Hospital of Asturias, Oviedo, Spain

^g Department of Neurology, University Clinical Hospital of Valladolid, Valladolid, Spain

^h Department of Neurology, University Hospital of Cruces, Barakaldo, Spain

ⁱ Health Research Institute Biocruces Bizkaia (Instituto de Investigación Sanitaria Biocruces Bizkaia), Barakaldo, Spain

^j Department of Radiology, Vall d'Hebron University Hospital, Barcelona, Spain

^k Multiple Sclerosis Division, Department of Neurology, La Princesa University Hospital, Madrid, Spain

^l Biomedical Research Institute of Málaga (Instituto de Investigación Biomédica de Málaga – IBIMA, Plataforma BIONAND), Málaga, Spain

^m Department of Neurology, Doctor Peset University Hospital, Valencia, Spain

ⁿ Department of Pharmacy, University Hospital Complex of A Coruña, A Coruña, Spain

^o Spanish Platform of Patients' Organisations (Plataforma de Organizaciones de Pacientes), Madrid, Spain

^p Department of Neurology, Río Carrión Hospital, Palencia, Spain

Received 12 July 2023; accepted 17 July 2023

KEYWORDS

Multiple sclerosis;
Clinical trial;
Critical appraisal

Abstract

Introduction: In recent years there has been an increase in the number of disease-modifying drugs (DMDs) approved for multiple sclerosis (MS). The evidence of their safety and efficacy has been obtained through several phase III and IV clinical trials. Acquiring the skills for their appraisal is indispensable for clinicians to assess the most pertinent treatment for patients. The objective of this study is to provide guidance in the critical reading of these trials.

* Corresponding author.

E-mail address: monica.perez.rios@usc.es (M. Pérez-Ríos).

PALABRAS CLAVE

Esclerosis múltiple;
Ensayo clínico;
Lectura crítica

Methods: A three-round e-Delphi study was carried out. In the preparatory phase, a multidisciplinary expert panel was established. Panel members were selected based on their scientific credentials and experience, seeking to include people involved in MS diagnosis, treatment and research. A semi-open questionnaire was developed based on key generic and MS-specific methodological instruments identified through a scoping bibliographic search. The experts were required to identify essential aspects for critically appraising clinical trials on DMDs for MS.

Results: The expert panel consisted of nine independent leading Spanish experts with long-standing experience with MS (five neurologists, a neuroradiologist, a pharmacologist, a research methodologist and an MS community representative). The e-Delphi study resulted in consensus recommendations intended to help readers in answering five major questions: "Is the study free of bias?"; "Are the included patients adequate?"; "Are the outcome measures appropriate?"; "Are the results relevant?"; and "Is the study transparent?".

Conclusion: This study proposes consensus recommendations intended to guide neurologists in the critical reading of phase III and IV clinical trials on DMDs for MS.

© 2025 Sociedad Española de Neurología. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Recomendaciones para la lectura crítica de ensayos clínicos sobre fármacos modificadores de la enfermedad para la esclerosis múltiple

Resumen

Introducción: En los últimos años ha habido un aumento en el número de fármacos modificadores de la enfermedad (FME) aprobados para la esclerosis múltiple (EM). La demostración de su seguridad y eficacia se ha logrado a través de diversos ensayos clínicos de fase III y IV. Adquirir las habilidades para su lectura crítica es indispensable para que los clínicos puedan determinar cuál es el tratamiento más pertinente para sus pacientes. El objetivo de este estudio es servir de guía para la lectura crítica de estos ensayos.

Métodos: Se llevó a cabo un estudio e-Delphi de 3 rondas. En la fase preparatoria se constituyó un panel multidisciplinario de expertos. Los miembros de este panel fueron seleccionados por su experiencia y credenciales científicas, buscando incluir personas implicadas en el diagnóstico, tratamiento e investigación en EM. A través de una búsqueda bibliográfica dirigida se identificaron instrumentos metodológicos genéricos y también específicos de EM, a partir de los cuales se elaboró un cuestionario semiabierto. En él se pidió a los expertos que se pronunciasen sobre qué aspectos son esenciales para evaluar críticamente ensayos clínicos sobre FME para la EM.

Resultados: El panel de expertos constó de 9 expertos españoles, independientes, y con larga experiencia en EM (5 neurólogos, un neurorradiólogo, un farmacólogo, un metodólogo de investigación y un representante de la comunidad). El e-Delphi se concretó en una serie de recomendaciones de consenso para ayudar a los lectores a responder 5 preguntas principales sobre cada ensayo: «¿está libre de sesgos?», «¿son adecuados los pacientes que se han incluido?», «¿son apropiadas las medidas de resultado utilizadas?», «¿son relevantes los resultados?» y «¿es transparente el estudio?».

Conclusión: Hemos elaborado una serie de recomendaciones de consenso destinadas a guiar a los neurólogos en la lectura crítica de los ensayos clínicos fase III y IV sobre FME para la EM.

© 2025 Sociedad Española de Neurología. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la CC BY-NC-ND licencia (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Multiple sclerosis (MS) is the most common demyelinating disease of the central nervous system and the leading non-traumatic cause of disability in young adults.¹ In 2020, approximately 2.8 million people were estimated to live with MS worldwide, with a global prevalence of 35.9 cases per 100,000 individuals, which is expected to increase in the coming years.² MS is more common in Europe and

North America, which both are areas of high economic development. This probably explains the advances in MS research that have occurred in the 21st century, leading to the approval within the last decade³ of numerous disease-modifying drugs (DMDs). DMDs aim to prevent relapses and ultimately intend to decrease the accrual of disability.⁴ These drugs alter the course of MS by suppressing or modulating the immune response involved in the disease pathogenesis but have potential risks of opportunistic infec-

tions, malignancies and other systemic adverse reactions. With the current therapeutic arsenal, and more DMDs to come in the near future, neurologists face challenges in recommending the most appropriate therapy for MS patients.

Clinical trials provide the rational basis for guiding neurologists and other physicians in their decision-making and clinical practice. However, their critical reading and analysis requires skills in deciding whether the available data provides credible evidence for the drug's safety and efficacy.⁵ Establishing the clinical benefit of a new drug requires an assessment of key issues related to the design and reporting of clinical trials, including the adequate selection of participants, the appropriateness of outcome measures, the adequacy of the statistical analysis, the clinical relevance of the results and the risk of bias. Without the appropriate skills, the physician's ethical and scientific obligation to reflect upon the adequacy of their clinical decisions⁶ would be replaced by a naive trust in evidence provided by third-parties.

The evaluation of clinical trials is especially complex in the case of MS, in which several concerns have been raised, relating to the appropriateness of the target population, the measurement of the progression of the disease, the selection of the appropriate active comparator or the choice of endpoints that would constitute a real clinical benefit, among others.⁷ For example, it has been suggested that the effectiveness of ocrelizumab on disability (the first approved drug for primary-progressive disease) was overestimated due to the inclusion of a young population and a considerable number of active (inflammatory) disease patients.⁸ Addressing these issues is especially problematic since there is no general consensus on many MS-specific methodological aspects nor recommendations for the assessment of these trials from the perspective of the neurologist.

The goal of this study is to elaborate a consensus document with key recommendations for the critical reading of phase III and IV clinical trials on DMDs for MS. This guidance is limited to pharmacological interventions in the adult population and excludes treatments for relapses and symptomatic drugs.

Materials and methods

A three-round e-Delphi study was carried out. A steering committee was convened to oversee its development, consisting of a neurologist, an epidemiologist and a public health physician.

Participant selection

A multidisciplinary expert panel was established by the steering committee using purpose sampling. The panel members were purposively selected based on their scientific credentials and experience, seeking to include participants with a minimum of 10 years of involvement in MS management and recognised leadership in the field. To account for diverse views within the MS research community, we attempted to include participants from different clinical and research backgrounds: clinicians currently working in a MS unit; researchers currently undertaking clinical or epidemi-

ological research on MS; decision-makers who have a current management role in MS; and representatives of MS patient organizations. Participants could simultaneously belong to more than one of these groups.

The experts were invited to participate in the study by means of a formal letter. They were required to identify essential aspects for the critical appraisal of clinical trials on MS. Emphasis was placed to remain anonymous, as well as to actively participate in all phases of the study to prevent their opinion from being underrepresented.

Preparatory round

A scoping bibliographic search was performed in PubMed to find key papers and methodological instruments for the proper design and critical reading of clinical trials, either generic or MS-specific. Additionally, panellists were invited to provide comments regarding the selected literature and propose additional relevant bibliographic material. Retrieved material was reviewed independently by two members of the steering committee and key information was extracted and rewritten in the form of simple statements that were converted into a questionnaire. Whenever possible, questionnaire items were grouped into different domains.

Rounds 1, 2 and 3

In round 1, experts were asked to individually rate the questionnaire items using a 1–9 Likert scale, whereby 1–3 indicates "limited importance", 4–6 "important but not essential" and 7–9 is "essential". Additionally, participants were encouraged to give constructive feedback (for that purpose, a comment box was provided alongside each statement) and suggest the addition of new items. Items rated as "limited importance" by at least 70% and "essential" by none of the respondents were rejected in this round. The rest of them were re-assessed in round 2.

In round 2, experts were asked to individually re-rate the questionnaire considering their own score in the previous round, as well as the comments and the median scores given by the panel. They were also asked to assess the new items that had been incorporated. Items that met the consensus criterion were included in the final document. Consensus was defined *a priori* as items rated as "essential" by at least 70% and of "limited importance" by less than 15% of the respondents.

Items from round 2 with a median score of 4–6 ("important but not essential") or 7–9 ("essential"), but not reaching the consensus criterion, were considered "dubious" and re-assessed in a third round. Items reaching consensus after this last round were also included in the final document.

In those cases where participants expressed difficulties in understanding any item, they were helped by the steering committee. When items were left unrated, the panellists were contacted for clarification. Special assistance was provided to the MS community representative to ensure a full understanding of the technical vocabulary used in the questionnaire.

Table 1 Expert panel characteristics (n = 9).

Sex, n (%)	
Male	6 (66.7)
Female	3 (33.3)
Age (years), median (IQR)	54 (50–59)
Professional experience (years), median (IQR)	20 (13–25)
Highest education background, n (%)	
MD/BSc/BA	3 (33.3)
Master	2 (22.2)
PhD	4 (44.5)
Area of expertise, n (%)	
Clinician	7 (77.8)
Researcher	8 (88.9)
Decision-maker	6 (66.7)
Patient representative	1 (11.1)
IQR: interquartile range.	

Data collection and analysis

The questionnaire was administered using the online survey platform Google Forms. A link to the questionnaire was sent to each of the panellists via e-mail. All experts were requested to complete each round within four weeks of receipt. If they did not return the survey, a reminder e-mail was sent. Anonymity was maintained throughout the process. Data was extracted and analysed using a predesigned Microsoft Excel worksheet.

Final document

Every item reaching consensus was included in a final document synthesising the panel's recommendations. Minor formatting changes were made, such as headline addition and content reordering.

Results

Expert panel

The expert panel consisted of nine leading Spanish experts with long-standing experience with MS: five neurologists, a neuroradiologist, a pharmacologist, a research methodologist and a MS community representative (Table 1). All of them formed part of leading national or international non-profit professional organizations and advisory groups involved in MS diagnosis, treatment or research. They participated in the study between August 2022 and January 2023. None of them received financial compensation.

Preparatory round

Through the scoping bibliographic search in PubMed, 16 key publications were identified (Table 2): the European Medicines Agency (EMA) "Guideline on clinical investigation of medicinal products for the treatment of Multiple Sclerosis",⁴ the Cochrane collaboration "Risk of Bias 2"

(RoB2) tool,⁹ the "Critical Appraisal Skills Programme" (CASP) checklist for clinical trials,¹⁰ the "Consolidated Standards of Reporting Trials (CONSORT) 2010" guideline¹¹ and twelve MS-specific publications.^{7,8,12–21} They were agreed upon and considered sufficient by the expert panel. After being reviewed, 86 key items were extracted and included in the questionnaire. They were grouped in five domains: internal validity, selection of participants, outcome measures, study results and transparency ([supplementary material](#)).

Rounds 1, 2 and 3

Participation rate was 100% in the three rounds. In round one, none of the 86 items was rejected and two new items were suggested by the panellists. In round 2, consensus was reached for 76 items. In round 3, consensus was reached for two additional items. Scores are detailed in [supplementary material](#).

Final document

The experts' consensus recommendations are shown in Table 3.

An overview of the whole e-Delphi is presented in Fig. 1.

Discussion

We have developed the first consensus recommendations to critically assess phase III and IV trials devoted to DMDs for MS. These recommendations, proposed by a multidisciplinary panel of national experts, can be easily applied by neurologists and guide them in the critical reading of these trials.

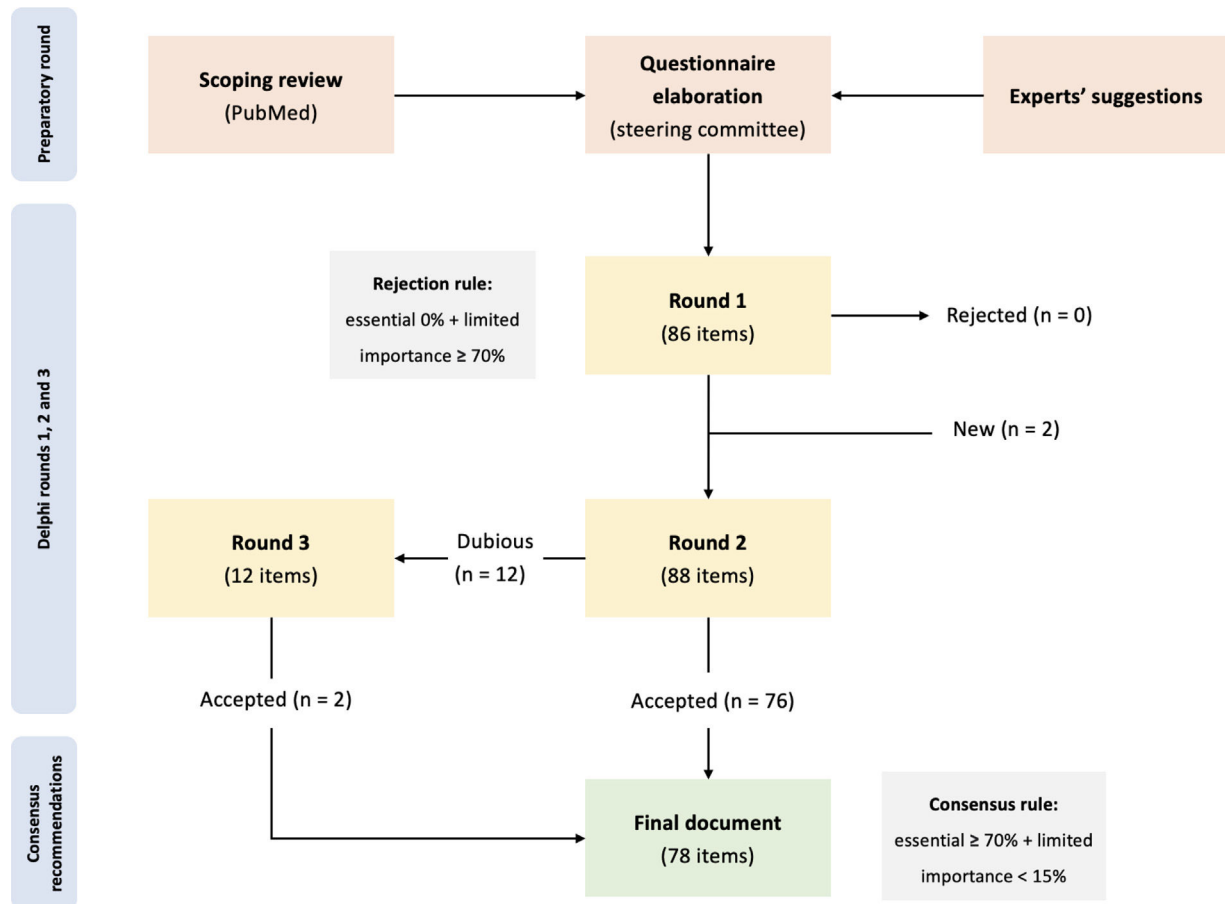
The field of MS research is constantly expanding and requires neurologists to update their knowledge in order to provide the most beneficial therapies.³ Despite constant bombardment of information about new DMDs, neurologists should be able to assess if a drug is safe and effective for a particular target population. To date, only generic critical reviews and proposals for the improvement of clinical trials in MS have been published.^{4,7,8,12–21} None were created from the reader's perspective and all of them leave relevant aspects unaddressed, such as the clinical relevance and applicability of the results. On the other hand, none of the generic tools developed to assess the adequate reporting of clinical trials (e.g. CONSORT statement) and to provide guidance for their critical reading (e.g. Critical Appraisal Skills Programme's checklist) take into account the specific aspects of MS clinical trials.

In this study, an e-Delphi exercise was carried out. The Delphi process is a widely used method for achieving consensus among experts on a given topic by means of an iterative, structured and transparent discussion process.²² It provides a method for collecting data based on the views of a panel of experts and, unlike group discussions, minimizes undue influences from domineering members. A key strength of the study is the recruitment of a multidisciplinary group of national experts with long-standing experience in MS

Table 2 Bibliography used for the elaboration of the questionnaire.

Author	Year	MS-specific	Document type	Scope
Brichetto and Zarin ¹⁴	2020	Yes	Review	Outcomes
Cochrane collaboration ⁹	2019	No	Assessment tool	Risk of bias
CONSORT group ¹¹	2010	No	Assessment tool	General
Critical Appraisal Skills Programme ¹⁰	2022	No	Assessment tool	General
European Medicines Agency ⁴	2015	Yes	Guideline	General
Gehr et al. ⁸	2019	Yes	Review	Outcomes
Krajnc et al. ¹⁵	2021	Yes	Review	Outcomes
Lublin et al. ¹⁶	2020	Yes	Review	Terminology
Marrie et al. ¹⁷	2016	Yes	Review	Comorbidity
Montalban ⁷	2011	Yes	Review	General
Ontaneda et al. ¹⁸	2015	Yes	Review	General
Pardini et al. ¹⁹	2019	Yes	Review	General
Tur et al. ²⁰	2018	Yes	Review	General
Uitdehaag ²¹	2018	Yes	Review	Outcomes
van Munster and Uitdehaag ¹²	2017	Yes	Review	Outcomes
Zhang et al. ¹³	2019	Yes	Review	General

MS: multiple sclerosis.

**Figure 1** e-Delphi summary.

and recognised leadership in different areas (including clinicians, researchers, decision-makers and patient representatives). While there is no agreement on the minimum sample size that ensures the accuracy of Delphi studies, it has been previously demonstrated that reliable results can

be obtained with small expert panels selected through strict criteria.²³ Furthermore, since the e-Delphi exercise was based on key items identified through a well-selected bibliography, we believe that the number of experts included was adequate and representative.

Table 3 Recommendations for the critical reading of phase III and IV clinical trials on disease-modifying drugs for multiple sclerosis.

Key question (dominion)	Aspects to be taken into account by the reader
I. Is the study free of bias? (internal validity)	1. The clinical trial should be randomized. 2. The clinical trial should be controlled with an active comparator or, in the event that this is not possible due to the absence of a therapeutic alternative, with placebo. If an active comparator is used, it should be of sufficient efficacy considering the patients' characteristics and their degree of disease activity. 3. The clinical trial should follow a parallel assignment. 4. Patients and healthcare, research and outcome assessment personnel should be blinded. 5. The clinical trial should follow a superiority design. 6. Losses to follow-up should not exceed 20% of the participants. The number of individuals who do not complete the study, their baseline characteristics and their reasons for dropping out should be similar between groups.
II. Are the included patients adequate? (selection of participants)	1. The clinical trial should use the MS diagnostic criteria that were valid at the time the study started. 2. In clinical trials evaluating the effect of the drug on relapses, patients with clinical activity and in early stages of the disease should be included. 3. In clinical trials evaluating the effect of the drug on PIRA, patients with evidence of sustained progression in the last two years should be included. 4. The characteristics of the included subjects should be similar to those of the patients who would receive the drug in normal clinical practice. 5. In order to assess the adequacy of the participants, the following baseline characteristics should be considered: - Sociodemographic: age, sex and ethnicity. - Clinical: EDSS, number of previous relapses*, number of recent relapses*, time since first symptom of MS, time since diagnosis of MS, time since conversion to secondary progressive MS**, previous use of DMDs and personal history of cancer*** or autoimmune thyroid diseases***. - Radiological (assessed by MRI): number of gadolinium-enhancing demyelinating lesions (T1)***, total number of demyelinating lesions (T2 or FLAIR), total volume of demyelinating lesions (T2 or FLAIR) and normalized brain volume.
III. Are the outcome measures appropriate? (outcome measures)	1. The primary outcome should be hard (i.e. clinically relevant) and not a surrogate or paraclinical endpoint. 2. In efficacy studies, the primary outcome should refer to either relapses or to accrual of disability (the latter being either due to PIRA or to RAW). In case of choosing one of the two previous outcomes as primary, the other one should be evaluated as key secondary. 3. To evaluate relapses, either the annualized relapse rate or the time to relapse should be used. In the case of time to relapse, a secondary outcome should be added to confirm the durability of the effect (e.g. time to a second or third relapse). 4. To assess accrual of disability, either the proportion of individuals who have worsened in a period of time or the time to that worsening should be used. Accrual of disability should not be defined as a mere increase in baseline disability, but must be pre-specified (e.g. the achievement of a certain degree of disability or a sustained worsening of a relevant magnitude) and confirmed, at least, after 6 months. The EDSS should be used, either alone or in combination with other clinical scales that increase its sensitivity to change (i.e. using a composite outcome). 5. Other non-primary outcomes should be assessed (when possible): - Clinical outcomes: cognition (SDMT and BICAMS are recommended), gait**** (T25FW and 6MWT are recommended), manual dexterity (9-HPT is recommended), visual function (LCLA is recommended), quality of life and patient-reported outcomes. - Radiological outcomes (assessed by MRI): number of new demyelinating lesions (T2 or FLAIR), number of gadolinium-enhancing demyelinating lesions (T1) and brain atrophy. - Other outcomes: NEDA-3*, NEPAD** and blood biomarkers (e.g. neurofilaments).

Table 3 Table 3 (Continued)

Key question (dominion)	Aspects to be taken into account by the reader
IV. Are the results relevant? (study results)	1. The clinical trial duration should be at least 2 years. 2. An intention-to-treat statistical analysis should be performed. 3. The efficacy results should be fully reported, indicating the intervention's effect size, confidence interval and <i>p</i>-value. 4. Efficacy should be stratified according to the disease subtype (relapsing-remitting, secondary progressive, primary progressive) and the degree of clinical and radiological activity (in the case of trials accepting patients with different phenotypes). 5. If the primary outcome result is statistically significant, the study should include an assessment of its clinical relevance and the risk–benefit ratio of the intervention. 6. Adverse effects should be detailed according to their severity and type. At least the following should be explicitly considered: neurological, psychiatric, infectious, autoimmune, neoplastic and cardiological.
V. Is the study transparent? (transparency)	1. The clinical trial should have been included in a public registry (e.g. ClinicalTrials.gov). This inclusion should have been done prospectively (i.e. before the inclusion of the first participant). 2. The clinical trial protocol should be publicly available. It should include the version history (i.e. the modifications that were made once the first patient was included), in which reasons for the changes ought to be stated. 3. The study should state whether or not it was sponsored by the pharmaceutical industry or any other third party. Each author should declare their conflicts of interest.

6MWT: six-minute walk test; 9-HPT: nine-hole peg test; BICAMS: brief international cognitive assessment for multiple sclerosis; DMDs: disease-modifying drugs; EDSS: expanded disability status scale; FLAIR: fluid attenuated inversion recovery; LCLA: low-contrast letter acuity; MRI: magnetic resonance imaging; MS: multiple sclerosis; NEDA-3: no evidence of disease activity 3; NEPAD: no evidence of progression or active disease; PIRA: progression independent of relapse activity; RAW: relapse-associated worsening; SDMT: symbol digit modalities test; T25FW: timed 25-foot walk.

* In relapsing forms.

** In progressive forms.

*** Unless exclusion criterion.

**** Except in non-ambulatory patients.

The Delphi questionnaire provided to the experts was developed from 16 relevant publications identified through a scoping bibliographic search. Although a systematic review was not undertaken, we consider that the possibility of missing relevant articles is minimal, since the information was cross-checked with experts. The most relevant document was the European Medicines Agency (EMA) "Guideline on clinical investigation of medicinal products for the treatment of Multiple Sclerosis",⁴ first published in May 2015. This document makes a series of proposals about how clinical trials on MS should be carried out, with special emphasis on the trial duration, the selection of adequate primary outcomes and the need to choose patients with appropriate phenotypes. However, it was written from the perspective of the clinical trial designer, not from the physician's. Furthermore, it has not been updated since 2015, it uses out-of-date terminology and it does not address relevant aspects such as the risk of biases and study relevance, applicability and transparency. The inclusion of the RoB2 tool,⁹ the CASP checklist for clinical trials,¹⁰ the CONSORT 2010 guidelines¹¹ and MS-specific updated bibliography allowed us to overcome these weaknesses when elaborating the questionnaire. The inclusion of all these documents gives a comprehensive approach to the consensus recommendations. In addition, grouping these recommendations into five categories allowed the correct assessment of five major questions: "Is the study free of bias?"; "Are the included

patients adequate?"; "Are the outcome measures appropriate?"; "Are the results relevant?"; and, finally, "Is the study transparent?".

As a result of the e-Delphi, the panel of experts reached a consensus on what key aspects should be considered by the readers when judging the reliability and quality of any clinical trials on DMDs for MS. In terms of a study's internal validity, the panel highlighted the need to use randomization, quadruple blinding and an active comparator (whenever possible), as well as avoiding attrition bias. Regarding the selection of participants, the experts proposed a core set of baseline characteristics (sociodemographic, clinical and radiological) that should be well specified so that the reader can determine whether the subjects included in the study are similar to those in clinical practice. Regarding outcome measures, the panel defended the use of hard primary outcomes (referring to either relapses or disability) and stated how to adequately assess accrual of disability. Likewise, it proposed a core set of non-primary outcomes along with the most adequate scales for their measurement. Regarding clinical trial results, the panel considered that they should be fully reported and stratified according to the MS subtype and degree of activity. In addition, it advocated for the trial promoter to reflect on the clinical relevance of these results. Finally, the panel emphasized the need for the study to be transparent, including its prospective inclusion in a public registry.

It is noteworthy that, throughout the e-Delphi, the participation rate was 100%. The steering committee did not consider it necessary to hold a face-to-face meeting due to the high degree of agreement obtained. This high agreement highlights the adequacy of the items included in the questionnaire. In addition, it must be taken into account that the aim of the panel was to make recommendations on critical aspects and that these do not constitute the definitive list of elements to be evaluated when critically reading a clinical trial on MS.

Conclusion

Choosing the appropriate treatment for MS patients is a complex decision that relies on deciding whether clinical trials offer credible evidence about a drug's safety and efficacy. In our study, we propose specific recommendations for neurologists to critically assess clinical trials on DMDs. This guidance is based on a transparent, well-known and easy-to-reproduce methodology and its availability could serve not only for critical reading, but also to promote a better design of future clinical trials involving MS patients.

CRedit authorship contribution statement

AR-d-A: study conceptualization, bibliographic search, data extraction, statistical analysis, interpretation of results, writing of the manuscript draft. MM-G: bibliographic search, interpretation of results, revision and edition of the manuscript. MP-R: methodology, revision and edition of the manuscript. MAL-G: Delphi panellist, revision and edition of the manuscript. PM: Delphi panellist, revision and edition of the manuscript. MM: Delphi panellist, revision and edition of the manuscript. ÀR: Delphi panellist, revision and edition of the manuscript. VM-L: Delphi panellist, revision and edition of the manuscript. AJG-R: Delphi panellist, revision and edition of the manuscript. LL: Delphi panellist, revision and edition of the manuscript. JRVH: Delphi panellist, revision and edition of the manuscript. PCR: Delphi panellist, revision and edition of the manuscript. AR-R: methodology, revision and edition of the manuscript. MP-H: interpretation of results, revision and edition of the manuscript. LV-L: study conceptualization, methodology, bibliographic search, interpretation of results, revision and edition of the manuscript, supervision of the manuscript. AR-d-A is responsible for the overall content as guarantor.

Funding

None declared.

Conflicts of interest

AR-d-A declares no conflicts of interest. MM-G declares no conflicts of interest. MP-R declares no conflicts of interest. MAL-G declares no conflicts of interest. PM declares no con-

flicts of interest. MM declares no conflicts of interest. ÀR serves or has served on the scientific advisory boards of BMS, Novartis, Sanofi-Genzyme, Synthetic MR, TensorMedical, Roche, Biogen and OLEA Medical, and has received financial support as a speaker in scientific activities from Bayer, Sanofi-Genzyme, Merck-Serono, Teva Pharmaceutical Industries Ltd, Novartis, Roche, BMS and Biogen. VM-L declares no conflicts of interest. AJG-R declares no conflicts of interest. LL declares no conflicts of interest. JRVH declares no conflicts of interest. PCR declares no conflicts of interest. AR-R declares no conflicts of interest. MP-H declares no conflicts of interest. LV-L declares no conflicts of interest.

Data availability

The study results are fully available.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version available at <https://doi.org/10.1016/j.nrleng.2025.08.003>.

References

- Filippi M, Bar-Or A, Piehl F, Preziosa P, Solari A, Vukusic S, et al. Multiple sclerosis. *Nat Rev Dis Prim*. 2018;4:43.
- Walton C, King R, Rechtman L, Kaye W, Leray E, Marrie RA, et al. Rising prevalence of multiple sclerosis worldwide: insights from the Atlas of MS, third edition. *Mult Scler J*. 2020;26:1816–21.
- Piehl F. Current and emerging disease-modulatory therapies and treatment targets for multiple sclerosis. *J Intern Med*. 2021;289:771–91.
- Committee for Medicinal Products for Human Use (CHMP), European Medicines Agency. Guideline on clinical investigation of medicinal products for the treatment of multiple sclerosis – Revision 2; 2015.
- du Prel J-B, Röhrig B, Blettner M. Critical appraisal of scientific articles. *Dtsch Arztebl International*; 2009.
- World Medical Association (WMA). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310:2191.
- Montalban X. Review of methodological issues of clinical trials in multiple sclerosis. *J Neurol Sci*. 2011;311:S35–42.
- Gehr S, Kaiser T, Kreutz R, Ludwig W-D, Paul F. Suggestions for improving the design of clinical trials in multiple sclerosis – results of a systematic analysis of completed phase III trials. *EPMA J*. 2019;10:425–36.
- Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;l4898.
- Critical Assessment Skills Programme. CASP Randomised Controlled Trial Standard Checklist [Internet]. Available from: <https://casp-uk.net/casp-tools-checklists/randomised-controlled-trial-rct-checklist/> [cited 26.02.23].
- Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340, c332.
- van Munster CEP, Uitdehaag BMJ. Outcome measures in clinical trials for multiple sclerosis. *CNS Drugs*. 2017;31:217–36.

13. Zhang Y, Salter A, Wallström E, Cutter G, Stüve O. Evolution of clinical trials in multiple sclerosis. *Ther Adv Neurol Disord*. 2019;12, 175628641982654.
14. Brichetto G, Zaratin P. Measuring outcomes that matter most to people with multiple sclerosis: the role of patient-reported outcomes. *Curr Opin Neurol*. 2020;33:295–9.
15. Krajnc N, Berger T, Bsteh G. Measuring treatment response in progressive multiple sclerosis – considerations for adapting to an era of multiple treatment options. *Biomolecules*. 2021;11:1342.
16. Lublin FD, Coetzee T, Cohen JA, Marrie RA, Thompson AJ. The 2013 clinical course descriptors for multiple sclerosis. *Neurology*. 2020;94:1088–92.
17. Marrie RA, Miller A, Sormani MP, Thompson A, Waubant E, Trojano M, et al. The challenge of comorbidity in clinical trials for multiple sclerosis. *Neurology*. 2016;86:1437–45.
18. Ontaneda D, Fox RJ, Chataway J. Clinical trials in progressive multiple sclerosis: lessons learned and future perspectives. *Lancet Neurol*. 2015;14:208–23.
19. Pardini M, Cutter G, Sormani MP. Multiple sclerosis: clinical trial design 2019. *Curr Opin Neurol*. 2019;32:358–64.
20. Tur C, Moccia M, Barkhof F, Chataway J, Sastre-Garriga J, Thompson AJ, et al. Assessing treatment outcomes in multiple sclerosis trials and in the clinical setting. *Nat Rev Neurol*. 2018;14:75–93.
21. Uitdehaag BMJ. Disability outcome measures in phase III clinical trials in multiple sclerosis. *CNS Drugs*. 2018;32:543–58.
22. Spranger J, Homberg A, Sonnberger M, Niederberger M. Reporting guidelines for Delphi techniques in health sciences: a methodological review. *Z Evid Fortbild Qual Gesundheitswes*. 2022;172:1–11.
23. Akins RB, Tolson H, Cole BR. Stability of response characteristics of a Delphi panel: application of bootstrap data expansion. *BMC Med Res Methodol*. 2005;5:37.