

Does the Multiple Sclerosis Impact Scale-29 (MSIS-29) have the range to capture the experience of fully ambulatory multiple sclerosis patients? Learnings from the ASCLEPIOS studies

Antoine Regnault, Angely Loubert, Róisín Brennan, Juliette Meunier, Christel Naujoks, Stefan Cano and Nicholas Adlard

Multiple Sclerosis Journal—
Experimental, Translational
and Clinical

July–September 2023, 1–12

DOI: 10.1177/
20552173231201422

© The Author(s), 2023.
Article reuse guidelines:
sagepub.com/journals-
permissions

Abstract

Background: Trials of disease-modifying therapies (DMTs) for multiple sclerosis (MS) often include patients with minimal disability. Patient-reported outcome instruments used in these trials have often not captured physical and psychological treatment effects concomitant with observed clinical benefits.

Objective: To examine whether the Multiple Sclerosis Impact Scale-29 (MSIS-29) captures changes in the impact of MS in a sample of patients enrolled in the Phase 3 ASCLEPIOS studies (ofatumumab vs. teriflunomide).

Methods: Measurement properties (i.e. item fit, reliability, and targeting) of the MSIS-29 were analyzed using Rasch measurement theory (RMT) in data from two phase 3 ofatumumab clinical trials including patients with relapsing-remitting or secondary progressive MS ($N=1882$). Targeting of the MSIS-29 items to the patient population was explored within groups categorized by Expanded Disability Status Scale (EDSS) scores.

Results: Under RMT analyses, both the Physical and Psychological Impact scales of the MSIS-29 were not appropriately targeted to the overall sample of patients. In particular, 49% and 30% of patients with an EDSS score ≤ 2.5 had fewer physical and psychological impacts, respectively, than would typically be captured by these MSIS-29 items compared to patients with EDSS scores of ≥ 3 .

Conclusion: The MSIS-29 is commonly used to evaluate the patient-reported physical and psychological impact of MS. However, it may be limited in evaluating changes associated with DMTs in patients with minimal disability.

Keywords: Disability, disease-modifying therapy, Multiple Sclerosis Impact Scale-29, multiple sclerosis, patient-reported outcomes, physical impact, psychological impact, psychometric analysis, Rasch measurement theory

Date received: 21 February 2023; accepted: 30 August 2023

Introduction

Multiple sclerosis (MS) is a chronic, immune-mediated disease of the central nervous system characterized by inflammation, demyelination and axonal/neuronal destruction, ultimately leading to severe disability.¹ Initiation of disease-modifying therapies (DMTs) earlier in the course of the disease is increasingly being considered to prevent disability

accrual.^{2–4} Therefore, new DMT clinical trials should use outcome measures sensitive to disability progression in recently diagnosed patients, who are likely to demonstrate limited functional impairments, reflected in lower Expanded Disability Status Scale (EDSS) scores. Disease progression in such patients may be episodic or tenuous and may not be detected by standard clinical outcome assessments included

Correspondence to:
Antoine Regnault,
Modus Outcomes,
4 Place Amedee Bonnet,
69002 Lyon, France.
antoine.regnault@
modusoutcomes.com

Antoine Regnault,
Angely Loubert,
Modus Outcomes, Lyon,
France



Róisín Brennan,
Novartis Pharma AG, Basel,
Switzerland

Juliette Meunier,
Modus Outcomes, Lyon,
France

Christel Naujoks,
Novartis Pharma AG, Basel,
Switzerland

Stefan Cano,
Modus Outcomes,
Letchworth Garden City, UK

Nicholas Adlard,
Novartis Pharma AG, Basel,
Switzerland

in clinical trials. These changes could be captured by patient-reported outcome (PRO) instruments that assess the impact of the disease on patients' daily lives. However, in clinical trials, DMTs have repeatedly failed to show benefits for MS on PRO endpoints.^{5–8} At best, as for cladribine⁹ or ocrelizumab¹⁰ clinical trials showed marginal positive impacts for some PRO domains. But the results were not confirmed in an independent study.¹¹ This could plausibly mean that either DMTs have no direct patient-reported physical and psychological benefits, or there is a limitation in the instrument used for measurement of impact in this patient population.

The Multiple Sclerosis Impact Scale-29 (MSIS-29) is a PRO instrument that has been used as an outcome measure in at least 34 MS clinical trials and was designed to assess the physical and psychological impact of MS.¹² MSIS-29 scores were collected in the phase 3 ASCLEPIOS I and II trials and it was observed that there was a discordance between the clinical and imaging outcomes, including disability progression, and the ability of the MSIS-29 instrument to accurately capture associated changes in the impact of MS.¹³

Several studies have evaluated the psychometric properties of the MSIS-29 in different clinical research settings.^{14–16} Most of the supportive evidence for MSIS-29 is based on traditional psychometric methods (classical test theory). It is now widely established that modern psychometric approaches (e.g. Rasch measurement theory [RMT]) provide a more systematic approach to appraising PRO instrument content.¹⁶ Applying RMT to MSIS-29 data highlighted some potential limitations, in particular the inability of MSIS-29 to adequately capture the impact of MS on patients with minimal disability.¹⁶ Cleanthous et al.¹⁷ recently explored the reconceptualization of the MSIS-29 scale using RMT analysis, which resulted in a three-scale structure: “Symptoms,” “General Limitations,” and “Psychological Impact.” The revised structure showed some improvement in the measurement properties of the MSIS-29. However, the poor coverage of patients with minimal disability levels was not fully resolved.

The present study examined the ability of the MSIS-29 to appropriately capture the impact of MS in a sample of patients with relapsing-remitting MS (RRMS) and secondary progressive MS (SPMS) from the ASCLEPIOS I/II phase 3 clinical trials, with a particular focus on the ability of the MSIS-29

to appropriately capture the impact of MS in patients with minimal disability (EDSS score ≤ 2.5 ; i.e., patients with mild disability in one of the seven functional systems (FS) or minimal disability in two).

Material and methods

Patient sample (ASCLEPIOS I and II studies)

Data from the randomized, double-blind, double-dummy, active comparator-controlled, parallel-group, multi-center, phase 3 ASCLEPIOS I and II studies investigating the efficacy, safety, and tolerability of subcutaneous ofatumumab versus oral teriflunomide in patients with relapsing forms of MS were used (detailed design and results of the ASCLEPIOS I and II studies have been published elsewhere¹⁸; for demographic and clinical characteristics at baseline, see Table 1). In brief, inclusion criteria consisted of adult relapsing MS patients aged between 18 and 55 years with active disease (RRMS or SPMS with an EDSS score of 0.5–5.5). Diagnosis of MS was defined according to the 2010 revised McDonald criteria¹⁹; RRMS was defined as recurrent acute exacerbations of neurologic dysfunction (relapses) followed by full or partial recovery; SPMS was defined as continuous worsening of disability that occurs independently of relapses.²⁰ Treatment duration varied for individual patients, with a maximal duration of 30 months (2.5 years) for each patient. The two studies were identically designed and conducted simultaneously, providing the rationale for using pooled data to conduct the psychometric analyses.

Quantitative analyses

The MSIS-29 was developed based on interviews of patients with RRMS, primary progressive MS, and SPMS, expert opinion, and literature review.¹² It includes 29 self-administered items originally grouped into two broad domains: Physical Impact (20 items) and Psychological Impact (nine items). Responses to questions about the impact of MS on different aspects of daily life are collected using a 4-point scale ranging from “not at all” (1) to “extremely” (4). The score obtained for each of the two scales by summing the responses to the individual items is linearly converted to a 0 to 100 range, where 100 indicates greater impact of MS.

In the ASCLEPIOS studies, the MSIS-29 questionnaire was collected at screening, baseline, and post-baseline visits at months 6, 12, 18, and 24. The MSIS-29 was also collected at the EOS visit and in case of treatment discontinuation (the end of treatment visit). Responses to the MSIS-29 items were

Table 1. Summary of the demographic and clinical characteristics (baseline visit, pooled data from the ASCLEPIOS I and II trials, N = 1882).

Parameter	Total N = 1882	EDSS ≤ 2.5 N = 932
Female, n (%)	1273 (67.6%)	628 (67.4%)
Male, n (%)	609 (32.4%)	304 (32.6%)
Age, mean (SD)	38 (9)	36 (9)
Age, n (%)		
18–30 years	442 (23.5%)	294 (31.5%)
31–40 years	651 (34.6%)	346 (37.1%)
41–55 years	785 (41.7%)	290 (31.1%)
> 55 years	4 (0.2%)	2 (0.2%)
Region, n (%)		
Eastern Europe	558 (29.6%)	258 (27.7%)
Western Europe	416 (22.1%)	235 (25.2%)
North America and Australia	426 (22.6%)	197 (21.1%)
Others	482 (25.6%)	242 (26.0%)
EDSS groups, n (%)		
Normal neurological examination (EDSS = 0)	44 (2.3%)	
Lower EDSS scores (1 ≤ EDSS score ≤ 2.5)	888 (47.2%)	
Intermediate EDSS scores (3 ≤ EDSS score ≤ 4.5)	747 (39.7%)	
Severe disability (5 ≤ EDSS score ≤ 5.5)	185 (9.8%)	
Very severe disability (EDSS score ≥ 6)	17 (0.9%)	
Missing	1 (0.1%)	
Type of MS at study entry, n (%)		
RRMS	1774 (94.3%)	930 (99.8%)
Active SPMS	108 (5.7%)	2 (0.2%)

EDSS: Expanded Disability Status Scale; MS: multiple sclerosis; RRMS: relapsing-remitting MS; SPMS: secondary progressive MS; SD: standard deviation.

described in subgroups of patients defined by their Expanded Disability Status Scale (EDSS) score.

RMT analyses were conducted with all visits combined (“stacked” data), without consideration for treatment groups. The MSIS-29 physical impact and psychological impact item sets were analyzed separately, as specified in the original instrument. RMT analyses were also performed using the alternative three-scale structure.¹⁶

Measurement properties from the RMT analysis, that is, fit to the Rasch model and reliability, were examined for each item set before focusing the interpretation on the targeting of the item set to the patient sample.¹⁶ Item fit was assessed based on ordering of item response options (i.e. ordering of item thresholds)²¹ and comparison of observed and expected responses using statistical indices (standardized fit residuals, using the range – 2.5 to +2.5 to characterize adequate fit,¹⁶ and chi-square tests) and graphical

examination of the item characteristic curve.²² Reliability was assessed using the Person separation index (PSI),²³ interpreted as follows: < 0.70, unsatisfactory; 0.70–0.79, modest; 0.80–0.89, adequate; 0.90–1.00, good reliability.²⁴ The targeting of the MSIS-29 item sets to the patient sample was examined by comparing the distribution of the estimates from the Rasch model for the patients and items. The main question of interest was whether the MSIS-29 item parameters adequately cover the part of the continuum corresponding to patients with a limited impact of MS.

Additionally, the distribution of the patient sample corresponding to the impact of MS resulting from the RMT analysis was described according to EDSS categories to illustrate how well the impact of MS was captured by the MSIS-29 items in patients with different levels of functioning, with a particular focus on patients with an EDSS score ≤ 2.5 (i.e. patients with mild disability in one of the seven FS

Table 2. Subgroups of patients defined by EDSS score.

EDSS groups n (%)	Total N (%)	Pooled visits N (%)
Normal neurological examination (EDSS = 0)	44 (2.3%)	266 (2.1%)
Lower EDSS scores ($1 \leq$ EDSS score ≤ 2.5)	888 (47.2%)	3927 (45.6%)
Intermediate EDSS scores ($3 \leq$ EDSS score ≤ 4.5)	747 (39.7%)	3027 (35%)
Severe disability ($5 \leq$ EDSS score ≤ 5.5)	185 (9.8%)	690 (8.0%)
Very severe disability (EDSS score ≥ 6)	17 (0.9%)	287 (3.3%)
Missing	1 (0.1%)	412 (4.8%)

Baseline visit. Total MS patients, N = 1882, and total pooled data points from two trials, N = 8609.
EDSS: Expanded Disability Status Scale.

or minimal disability in two). The distribution of patients reporting a given level of impact across each of the MSIS-29 items was also depicted graphically using heatmaps.

Analyses were performed using RUMM 2030 software (RUMM Laboratory, Perth, Australia) for RMT analysis and SAS v9.4 (SAS Institute Inc., Cary, NJ, USA) for data preparation and all other analyses.

Results

Sample description

Briefly, ASCLEPIOS I and II included a total of 1882 patients (mean age across both trials was 38 years, 68% were female [Table 1]). Most were from Europe (Eastern Europe, 30%; Western Europe, 22%) or from North America and Australia (23%). Most patients had RRMS (94%) and 6% had SPMS. Almost half of the patients (49.5%) had normal neurological examination or minimal-to-no disability (EDSS score ≤ 2.5) at baseline. The change from baseline in MSIS-29 scores for the ofatumumab and teriflunomide groups over the course of the pooled ASCLEPIOS I and II trials are presented in Supplemental Material 1.

Item responses at pooled visits were described within subgroups of patients defined by EDSS score (Table 2). These subgroups were used in the subsequent RMT analyses by EDSS score and heatmaps. A total of 8609 assessments of the MSIS-29 were performed over the course of the two studies and were included in the analyses.

Physical impact domain

Distribution of responses to MSIS-29 physical impact items. The distribution of responses to the MSIS-29 physical impact items according to EDSS score is

presented in Figure 1. In patients with EDSS score ≥ 3 , the response to most items typically indicated at least “a little” impact of MS. In patients with an EDSS score ≤ 2.5 , the most frequent response reported no impact for any of the physical impact items. The only items for which $> 40\%$ of patients with minimal disability levels (EDSS score ≤ 2.5) reported “a little” impact were “Doing physically demanding tasks,” “Problems with balance,” and “Heavy arms or legs.”

RMT analysis of MSIS-29 physical impact domain.

The RMT analysis of the MSIS-29 physical impact domain in the overall sample showed mixed results. All items had correctly ordered thresholds, indicating that the response categories worked as intended (Figure 2(b)) and the scale had good reliability (PSI = 0.91). However, examination of the item fit showed observed responses to these items were not always aligned with predicted model outcomes; in fact there were deviations from the Rasch model for most items (standardized fit residuals outside of the $-2.5/2.5$ range and significant chi-square statistics [Supplemental Material 2]).

When considering the coverage of the physical impact continuum, very few “item thresholds” were distributed at the lower, left-hand end of the continuum (Figure 2(b) and (c)). The items with the lowest estimated first “threshold” (corresponding to the shift between “no impact” and “a little impact”), indicating the mildest physical impact captured by the MSIS-29, were for being “a little limited” in “doing physically demanding tasks,” and being “a little bothered” by “problems with balance,” “heavy arms or legs,” “taking longer to do things,” or “being clumsy.” In contrast, a substantial proportion of the patient sample was distributed toward this end of the continuum (Figure 2(a)): $> 20\%$ of the patient sample had lower physical impact than would typically be

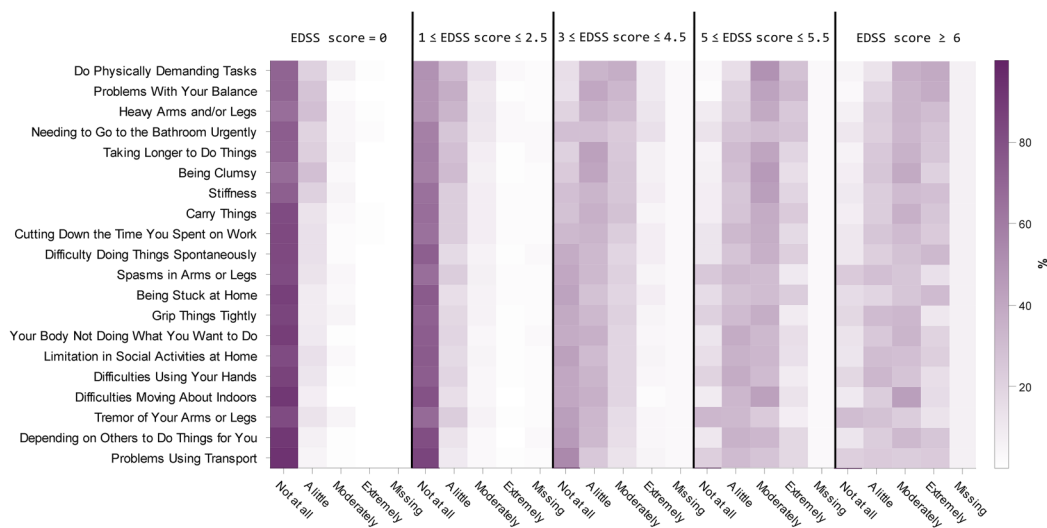


Figure 1. Heatmap of the responses to Multiple Sclerosis Impact Scale-29 (MSIS-29) “Physical Impact” domain items (pooled visits, pooled data from two trials, $N = 8197$). Each cell of the map shows the percentage of patients that reported a given level of impact (column) for a given item (row), with darker fill colors indicating higher percentages.

captured by any of the MSIS-29 items. As expected, the small number of items adequate for low levels of physical impact led to substantial uncertainty associated with measurement of this portion of the continuum (Figure 2(d)).

Distribution of MSIS-29 physical impact measures according to EDSS score. A plot of the distribution of the patient sample over the physical impact continuum from the RMT analysis in the subgroups of patients defined by baseline EDSS score was highly skewed toward very low levels of impact in patients with an EDSS score ≤ 2.5 (Figure 3). Almost half (49%) of patients with an EDSS score ≤ 2.5 had lower physical impact than would typically be captured by any of the MSIS-29 items. The physical impact of patients with EDSS scores of ≥ 3 aligned much better with the MSIS-29 items.

Psychological impact domain

Distribution of responses to MSIS-29 psychological impact items. Similar to the physical impact domain, a clear difference was observed in the distribution of responses to the psychological impact items depending on EDSS score (Figure 4). Patients with an EDSS score ≥ 3 typically reported at least “a little impact” for most psychological impact items, whereas the most frequent response reported by patients with EDSS score ≤ 2.5 was “no impact” for all psychological impact domain items. However, a higher proportion of patients with minimal disability

levels (EDSS score ≤ 2.5) reported some psychological impact across all items than was reported for physical impact items.

RMT analysis of MSIS-29 psychological impact domain. The RMT analysis of the MSIS-29 psychological impact domain in the overall sample showed mixed results. All items had correctly ordered thresholds, indicating that the response categories worked as intended (Figure 5(b)) and the scale had adequate reliability (PSI = 0.86). Similar to the physical impact analysis, observed responses to these items were not always aligned with predicted model outcomes and the item fit showed deviations from the Rasch model for most items (standardized fit residuals outside of the $-2.5/2.5$ range and significant chi-square statistics [Supplemental Material 3]).

Only a few psychological impact item thresholds were estimated at the lower end of the continuum (Figure 5(b) and (c)). The lowest estimated first thresholds (corresponding to the shift between “no impact” and a “little impact”) were for being “a little bothered” by “feeling mentally fatigued,” “being irritable,” “having problems concentrating,” “worries related to MS,” “feeling anxious,” or “feeling unwell.” Nonetheless, almost 30% of the patient sample had lower psychological impact than would be typically captured by any MSIS-29 item (Figure 5(a)). Again, and as expected, the small numbers of items adequate for low levels of psychological impact resulted in substantial uncertainty

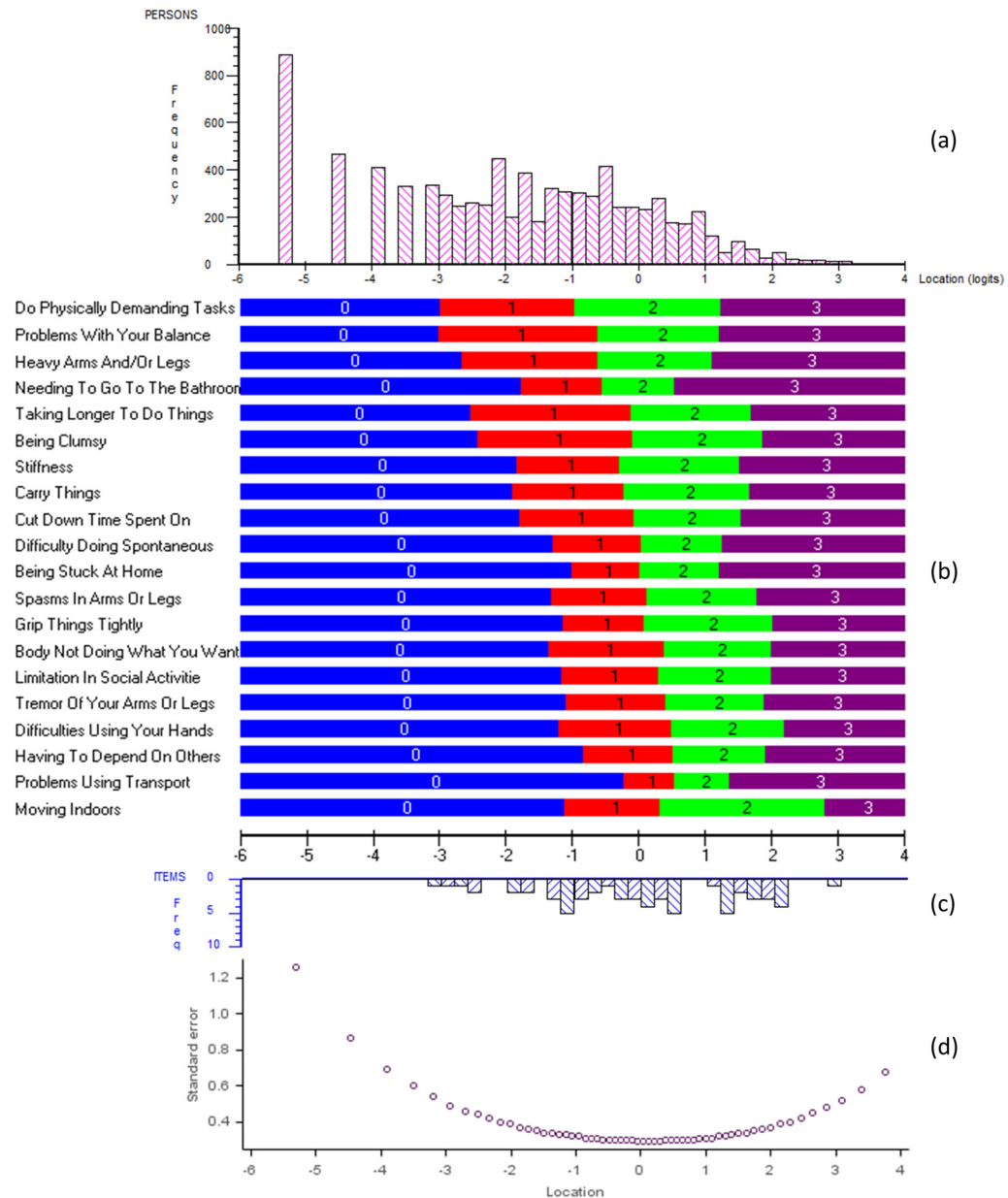


Figure 2. Results of Rasch measurement theory analysis of the Multiple Sclerosis Impact Scale-29 (MSIS-29) “Physical Impact” domain (pooled visits, pooled data from two trials, N = 8389): (a) distribution of estimated parameters for each individual MSIS-29 assessment over the continuum of physical impact; (b) MSIS-29 item threshold map, presenting the most probable response to each item along the continuum of physical impact; (c) distribution of estimated MSIS-29 item thresholds over the continuum of physical impact; (d) standard error associated with individual MSIS-29 assessment over the continuum of physical impact. The same X-axis is shared by all graphs to allow direct comparison between panels; and defines the continuum of psychological impact measured by the MSIS-29 domain as estimated by the Rasch model. It is expressed in “logits” (the unit of the Rasch model) and ranges from lower psychological impact on the left to higher psychological impact on the right.

associated with the measurement of this portion of the continuum (Figure 5(d)).

Distribution of MSIS-29 psychological impact measures according to EDSS score. Similar to results for the physical impact measures, there is a

clear mismatch between the psychological impact of patients with low EDSS scores (≤ 2.5) and MSIS-29 items (Figure 6); >40% of the subgroup of patients with an EDSS score ≤ 2.5 had lower psychological impact than would typically be captured by any of the MSIS-29 items. However, the

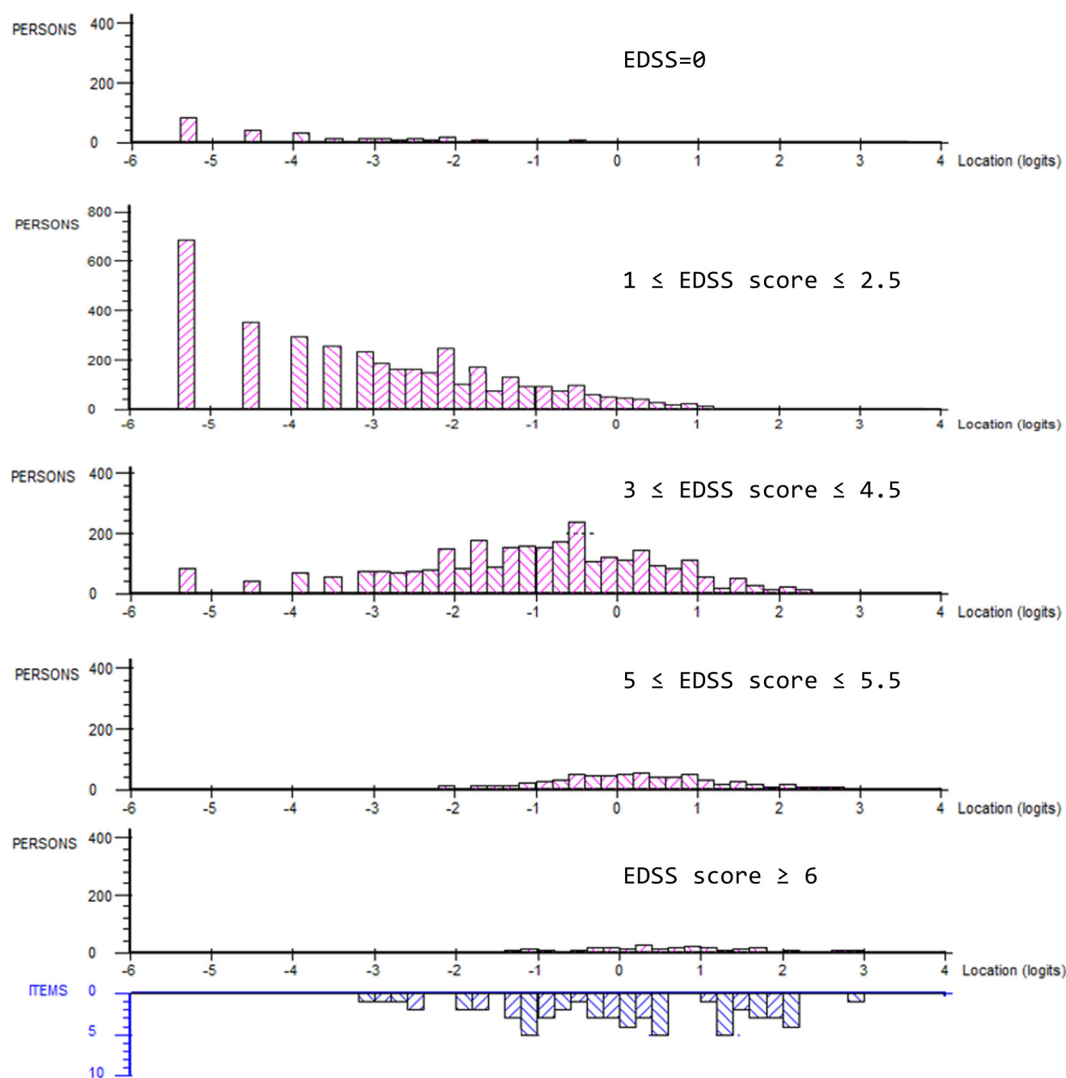


Figure 3. Distribution of estimated parameter for each MSIS-29 “Physical Impact” assessment within EDSS score (pink bars) and of estimated MSIS-29 item thresholds (blue bars) over the same metric of physical impact from the Rasch model (pooled visits, pooled data from two trials, $N = 7977$). The same X -axis is shared by all graphs to allow direct comparison between panels; and defines the continuum of psychological impact measured by the MSIS-29 domain as estimated by the Rasch model. It is expressed in “logits” (the unit of the Rasch model) and ranges from lower psychological impact on the left to higher psychological impact on the right.

MSIS-29: Multiple Sclerosis Impact Scale-29; EDSS, EDSS: Expanded Disability Status Scale.

psychological impact of patients with an EDSS scores ≥ 3 aligned much better with the MSIS-29 items.

corresponding to low impact, where a substantial proportion of the patient sample was distributed.

Alternative three-scale structure

The RMT analysis of the alternative three-scale structure (symptoms, psychological impact, and general limitations) did not show significantly improved results compared with the original structure in terms of coverage of the low-impact end of the measured continuums (Supplemental Materials 4 to 6). For all three modified scales, no MSIS-29 items were located on the portion of the continuum

Discussion

RMT analyses of the MSIS-29 data from ASCLEPIOS I and II highlighted the limitations of the MSIS-29 instrument to capture the impact of MS in minimally disabled patients. Only a few MSIS-29 items (i.e. being “a little bothered” not being able to perform physically demanding tasks, by troubles with balance, or by mental fatigue) appeared to adequately capture the physical or

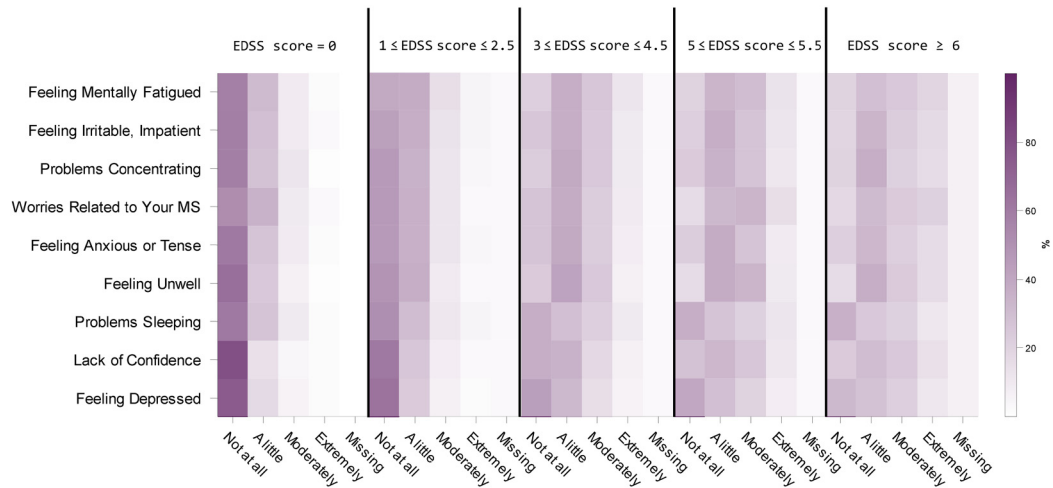


Figure 4. Heatmap of the responses to Multiple Sclerosis Impact Scale-29 (MSIS-29) “Psychological Impact” domain items (pooled visits, pooled data from two trials, N = 8197). Each cell of the map shows the percentage of patients that reported a given level of impact (column) for a given item (row), with darker fill colors indicating higher percentages.

psychological impact of MS relevant to patients with low EDSS scores.

Cleanthous et al.¹⁷ aimed to document the measurement properties of the MSIS-29 and explore possible different scoring of the MSIS-29. The analysis we present was similar as we also explore the measurement properties of the MSIS-29. However, the previous research did not explore the relationship of the MSIS-29 with the EDSS level, which is a central to our research, and therefore constitutes a clear added value to previous analyses

The Cleanhouse et al.¹⁷ analysis represents a previous application of the Rasch model that demonstrated mistargeting of MSIS-29 physical impact and psychological impact scales to an RRMS clinical trial sample. In ASCLEPIOS I and II, the mistargeting of the MSIS-29 items was primarily observed in subgroups of patients with minimal disability levels (EDSS scores ≤ 2.5). This confirmed the different manifestation of the physical and psychological impact of MS in these patients versus those with more debilitating MS. This difference may be in the degree (milder impact that cannot be captured by the questions as worded in the MSIS-29) or in the nature of the impact (different areas of daily life impacted or impact taking a different form). The measurement of the impact of MS by the MSIS-29 in patients with minimal disability levels is therefore probably incomplete and associated with much uncertainty. In this context, the assessment of the impact of MS in minimally disabled patients needs to be reconsidered.

An alternative domain structure for the MSIS-29 has been proposed¹⁷ with the intent of improving the psychometric performance. However, these modifications do not address the measurement of the impact of MS in patients with no-to-mild disability. It should be noted these data were not adjusted for sample size; therefore, the resulting chi-square should be interpreted in the context of a large sample size. Further research is required to gain a better understanding of the daily experience of patients with lower EDSS scores. It would be important to identify the specific areas of their lives impacted and the manifestation of this impact. This would allow the generation of a set of patient-reported items targeted to this subgroup of patients. These items could be used as a standalone measure targeting patients with MS and limited functional impairment. The items could also be used jointly with the MSIS-29, providing wider coverage of the impact of MS over different stages of the disease if a broader population is considered. Such a “bolt-on” strategy was recently applied to address similar issues in other PRO instruments used in MS.^{25,26}

Despite the superiority of ofatumumab over teriflunomide on reducing the risk of relapses and disability progression in ASCLEPIOS I/II—outcomes which directly affect quality of life—the MSIS-29 did not demonstrate a clinically meaningful difference between the treatment arms. Demonstration of a treatment benefit in this context requires instruments able to detect subtle changes in the experience of patients whose functional status is minimally impaired.

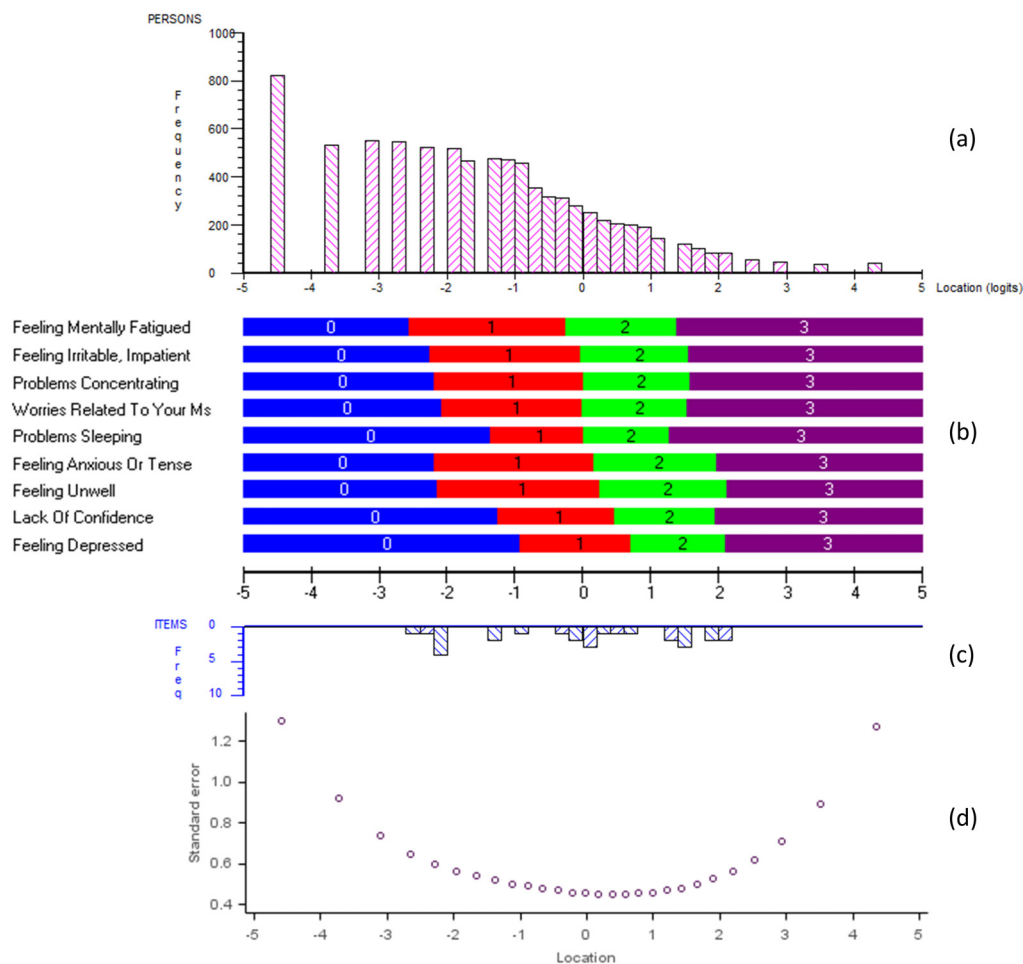


Figure 5. Results of Rasch measurement theory analysis of the Multiple Sclerosis Impact Scale-29 (MSIS-29) “Psychological Impact” domain (pooled visits, pooled data from two trials, $N = 8375$): (a) distribution of estimated parameters for each individual MSIS-29 assessment over the continuum of psychological impact; (b) MSIS-29 item threshold map, presenting the most probable response to each item along the continuum of psychological impact; (c) distribution of estimated MSIS-29 item thresholds over the continuum of psychological impact; (d) standard error associated with individual MSIS-29 assessment over the continuum of psychological impact. The same X-axis is shared by all graphs to allow direct comparison between panels; and defines the continuum of psychological impact measured by the MSIS-29 domain as estimated by the Rasch model. It is expressed in “logits” (the unit of the Rasch model) and ranges from lower psychological impact on the left to higher psychological impact on the right.

Our analyses of the MSIS-29 offer a possible explanation for the absence of conclusive PRO results in MS clinical trials. The lack of items capturing the impact of MS in patients experiencing lower disability levels can be explained by the development process of the MSIS-29. Historically, the items in the MSIS-29 final version were selected using data from a postal survey of a random sample of members of the Multiple Sclerosis Society of Great Britain and Northern Ireland around the year 2000.¹² Limited information is available to characterize this sample (e.g. no EDSS was collected). The inclusion criteria of the ASCLEPIOS trials were based on the McDonald criteria revised in 2010. It is therefore

reasonable to assume that the MSIS-29 development sample included mostly patients experiencing greater levels of disability. In such a sample, the items relevant to patients with the least functional impairment were likely not endorsed by the majority of the sample and may have shown lower correlations than the items more targeted to the majority of the sample. With this approach, it is likely that the items reflecting the least impact of MS were not included in the instrument. This approach was appropriate for the original intended use of MSIS-29 in a broad MS population under the diagnostic criteria of the time, but not fully for a population recruited under revised diagnostic criteria that emphasize the

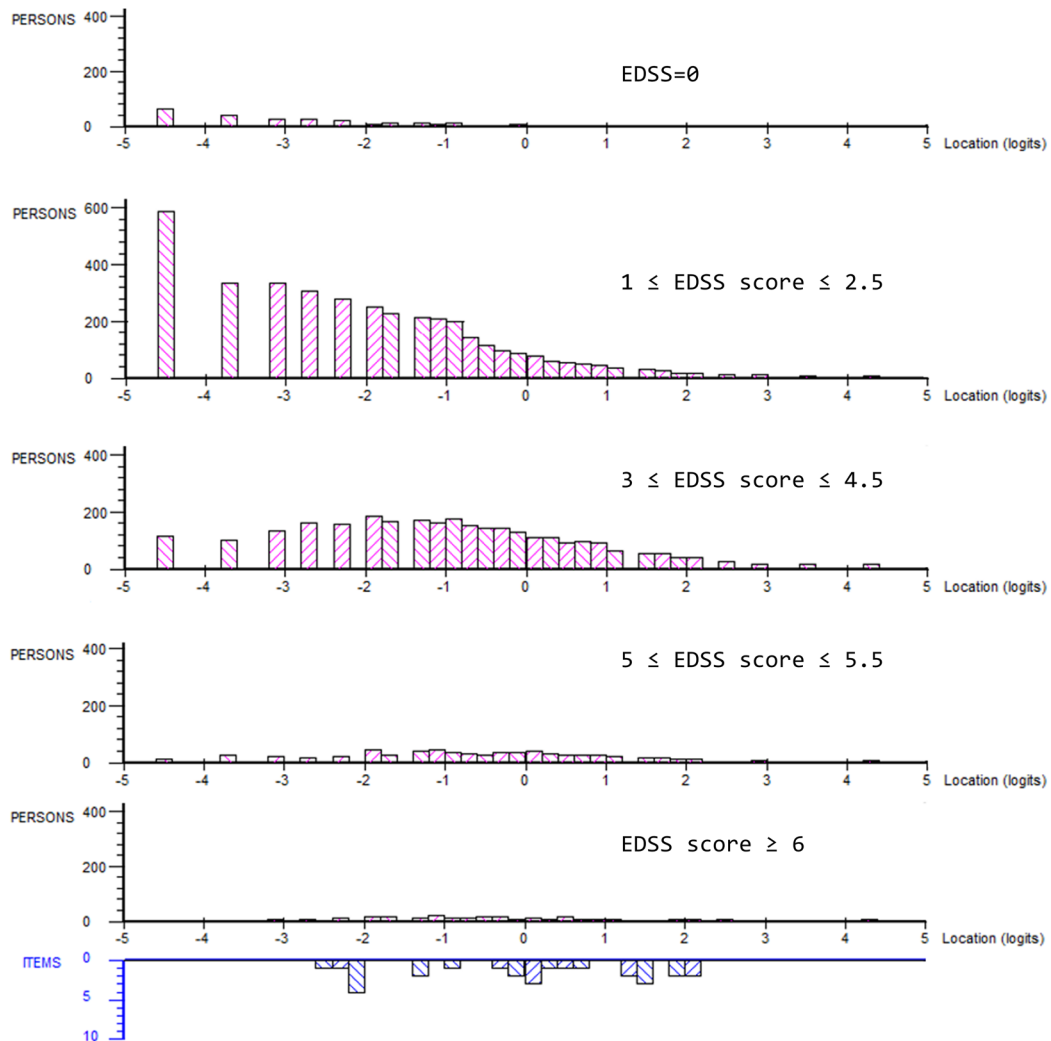


Figure 6. Distribution of estimated parameter for each MSIS-29 “Psychological impact” assessment within EDSS score (pink bars) and of estimated MSIS-29 item thresholds (blue bars) over the same metric of psychological impact from the Rasch model (pooled visits, pooled data from two trials, $N = 7963$). The same X -axis is shared by all graphs to allow direct comparison between panels; and defines the continuum of psychological impact measured by the MSIS-29 domain as estimated by the Rasch model. It is expressed in “logits”, the unit of the Rasch model and ranges from lower psychological impact on the left to higher psychological impact on the right.

MSIS-29: Multiple Sclerosis Impact Scale-29; EDSS, EDSS: Expanded Disability Status Scale.

need for early treatment. It is worth noting that our analysis is limited to cross-sectional data and that a future analysis on the longitudinal changes in the MSIS-29 in the ASCLEPOIS trial population may offer insights into the effect of noise on the outcome measure and strengthen the demonstration of the limitations of the MSIS-29.

In MS research, PROs are used as outcome measures,^{16,27,28} however it is worth noting that, similarly to the development of the MSIS-29, the majority of PROs used in MS research were developed within populations that are majorly different from the

average MS population. Moreover, due to progress in the MS treatment landscape, and the earlier prescription of DMTs, the course of clinical worsening in the majority of MS patients has changed since most PROs used in MS clinical trials were developed.²⁹ A PRO measure more aligned with recent diagnostic criteria and reflecting a paradigm centered on early treatment would provide a more accurate demonstration of the benefits of DMTs. In fact, the specific nature of DMTs warrants critically revisiting how their efficacy in terms of PROs should be investigated in MS clinical trials, going beyond the mere question of the outcome measure to be used. More

specifically, identifying well-defined PRO objectives, including the identification of the patient-relevant concepts that should be primarily targeted, is recommended as good practice for clinical trial protocol development,³⁰ guiding the specification of PRO estimands,³¹ and statistical analysis of PRO endpoints.³²

As further advances in the identification and treatment of MS are made, the conceptualization of the impact of MS in minimally disabled patients will be a fundamental step towards demonstrating successful PRO benefits in future DMT trials.

Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Antoine Regnault, Angély Loubert, Juliette Meunier, and Stefan Cano are employees of Modus Outcomes, which was commissioned by Novartis to conduct the work. Róisín Brennan, Christel Naujoks, and Nicholas Adlard are employees and shareholders of Novartis Pharma AG.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by Novartis Pharma AG, Basel, Switzerland.

Supplemental material

Supplemental material for this article is available online.

References

- Montalban X, Gold R, Thompson AJ, et al.ECTRIMS/EAN guideline on the pharmacological treatment of people with multiple sclerosis. *Mult Scler* 2018; 24: 96–120.
- Rae-Grant A, Day GS, Marrie RA, 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: 777–788.
- Brown JW, 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.
- Fernandez O. Is there a change of paradigm towards more effective treatment early in the course of apparent high-risk MS? *Mult Scler Relat Disord* 2017; 17: 75–83.
- Calabresi PA, Radue EW, Goodin D, et al. Safety and efficacy of fingolimod in patients with relapsing-remitting multiple sclerosis (FREEDOMS II): a double-blind, randomised, placebo-controlled, phase 3 trial. *Lancet Neurol* 2014; 13: 545–556.
- Kapoor R, Ho PR, Campbell N, et al. Effect of natalizumab on disease progression in secondary progressive multiple sclerosis (ASCEND): a phase 3, randomised, double-blind, placebo-controlled trial with an open-label extension. *Lancet Neurol* 2018; 17: 405–415.
- Kita M, Fox RJ, Phillips JT, et al. Effects of BG-12 (dimethyl fumarate) on health-related quality of life in patients with relapsing-remitting multiple sclerosis: findings from the CONFIRM study. *Mult Scler* 2014; 20: 253–257.
- (ON) O. *Teriflunomide (Aubagio) (14 mg film-coated tablet): Teriflunomide is indicated as monotherapy for the treatment of patients with relapsing-remitting multiple sclerosis to reduce the frequency of clinical exacerbations and to delay the accumulation of physical disability*. Ottawa, ON: Canadian Agency for Drugs and Technologies in Health; 2014 Oct. CADTH Common Drug Reviews, 2014.
- Afolabi D, Albor C, Zalewski L, et al. Positive impact of cladribine on quality of life in people with relapsing multiple sclerosis. *Mult Scler* 2018; 24: 1461–1468.
- Hauser SL, Bar-Or A, Comi G, et al. Ocrelizumab versus interferon beta-1a in relapsing multiple sclerosis. *N Engl J Med* 2017; 376: 221–234.
- Zhu C, Zhou Z, Roos I, et al. Comparing switch to ocrelizumab, cladribine or natalizumab after fingolimod treatment cessation in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2022; 93: 1330–1337.
- Hobart J, Lamping D, Fitzpatrick R, et al. The multiple sclerosis impact scale (MSIS-29): a new patient-based outcome measure. *Brain* 2001; 124: 962–973.
- Health OOCaFdaTi. *Clinical review report: Ofatumumab (Kesimpta)*. Ottawa, ON, Canada: Novartis Pharmaceuticals Canada Inc., 2021.
- Khurana V, Sharma H, Afroz N, et al. Patient-reported outcomes in multiple sclerosis: a systematic comparison of available measures. *Eur J Neurol* 2017; 24: 1099–1107.
- I. ADaM. *A course in Rasch measurement theory*, 2019, p.41–53.
- Hobart J and Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. *Health Technol Assess* 2009; 13: iii, ix-x, 1-177. DOI: 10.3310/hta13120.
- Cleanthous S, Cano S, Kinter E, et al. Measuring the impact of multiple sclerosis: enhancing the measurement performance of the Multiple Sclerosis Impact Scale (MSIS-29) using Rasch measurement theory (RMT). *Mult Scler J Exp Transl Clin* 2017; 3: 2055217317725917.
- Hauser SL, Bar-Or A, Cohen JA, et al. Ofatumumab versus teriflunomide in multiple sclerosis. *N Engl J Med* 2020; 383: 546–557.
- Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol* 2011; 69: 292–302.

20. Lublin FD, Reingold SC, Cohen JA, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology* 2014; 83: 278–286.
21. Andrich D. Rating scales and Rasch measurement. *Expert Rev Pharmacoecon Outcomes Res* 2011; 11: 571–585.
22. GN WBaM. *Rating scale analysis*. Chicago, IL: Mesa Press, 1982.
23. Andrich D. An index of person separation in latent trait theory, the traditional KR. 20 index, and the Guttman scale response pattern. *Education Research and Perspectives* 1982; 9: 95–104.
24. IH NJaB. *Psychometric theory*, 1994.
25. Strzok S, Cleanthous S, Pompilus F, et al. Development of a gait module to complement the 12-item Multiple Sclerosis Walking Scale: a mixed methods study. *Mult Scler J Exp Transl Clin* 2018; 4: 2055217318783766.
26. Cleanthous S, Strzok S, Pompilus F, et al. Addressing the targeting range of the ABILHAND-56 in relapsing-remitting multiple sclerosis: a mixed methods psychometric study. *Mult Scler J Exp Transl Clin* 2018; 4: 2055217318776990.
27. Hobart J, Chitnis T, Oh J, et al. Do clinical trials prepare to fail by failing to prepare? An examination of MS trials and recommendations for patient-reported outcome measure selection. *Mult Scler Relat Disord* 2023; 76: 104788.
28. Hobart JC, Cano SJ, Zajicek JP, et al. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations. *Lancet Neurol* 2007; 6: 1094–1105.
29. Granziera C, Derfuss T and Kappos L. Time to change the current clinical classification of multiple sclerosis? *JAMA Neurol* 2023; 80: 128–130.
30. Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: the SPIRIT-PRO extension. *JAMA* 2018; 319: 483–494.
31. Fiero MH, Pe M, Weinstock C, et al. Demystifying the estimand framework: a case study using patient-reported outcomes in oncology. *Lancet Oncol* 2020; 21: e488–e494.
32. Coens C, Pe M, Dueck AC, et al. International standards for the analysis of quality-of-life and patient-reported outcome endpoints in cancer randomised controlled trials: recommendations of the SISAQOL consortium. *Lancet Oncol* 2020; 21: e83–e96.