

# A comparison of clinical outcomes in PPMS in the INFORMS original trial data set

Marcus W Koch , Jop P Mostert, Bernard Uitdehaag and Gary Cutter

## Abstract

**Background:** The expanded disability status scale (EDSS) is the standard clinical outcome measure in primary progressive multiple sclerosis (PPMS), even though the timed 25-foot walk (T25FW), nine-hole peg test (NHPT) or combinations of these measures may be more useful. The paced auditory serial addition test (PASAT) is a widely used cognitive measure in MS, but little is known about change in PASAT scores over time in PPMS.

**Objective:** The objective of this study is to compare clinical outcome measures in a large PPMS trial data set.

**Methods:** We determined significant worsening events on the EDSS, T25FW and NHPT, and PASAT scores over the course of this 3-year trial. We compared unconfirmed, confirmed and sustained disability worsening and contrasted disability worsening with similarly defined improvement. We examined the association of baseline characteristics with the risk of disability worsening at 12, 24 and 36 months with logistic regression models.

**Results:** The EDSS and T25FW showed most worsening events, while only few patients worsened on the NHPT. Adding the NHPT to a combined outcome added only few further worsening events. PASAT scores slightly increased over time, possibly due to a practice effect.

**Conclusion:** Both the EDSS and T25FW, but not NHPT or PASAT, appear to be useful outcome measures in PPMS.

**Keywords:** Primary progressive multiple sclerosis, clinical trials methodology, outcome measures

Date received: 26 October 2020; revised: 20 December 2020; accepted: 22 December 2020.

## Introduction

The treatment of primary progressive multiple sclerosis (PPMS) remains an unmet challenge. To date, only one treatment for PPMS is available, and this treatment only has a modest effect, predominantly in people with remaining focal inflammatory disease activity.<sup>1</sup> Meaningful progress in the development of new and impactful treatments for PPMS will undoubtedly require many more clinical trials investigating new interventions. However, clinical trials in PPMS are difficult and expensive to conduct, in part due to the use of the expanded disability status scale (EDSS)<sup>2</sup> as the primary outcome measure in trials in all forms of MS.<sup>3</sup>

The use of the EDSS in trials has become the standard with a time to event outcome of confirmed disability progression (CDP). Trials using this outcome are

powered based on the number of progression events, which occur between 30% and 40% of participants over the course of 2 years. Using this approach translates into large sample sizes. If alternative outcome measures have higher event rates, it might be possible to lower the sample size and shorten the duration of trials.

We recently investigated clinical outcome measures in two secondary progressive MS (SPMS) trial data sets and found that the timed 25-foot walk (T25FW)<sup>4</sup> may be a more useful primary outcome measure than the EDSS. For this study, we gained access to the data set of the INFORMS trial,<sup>5</sup> a large phase III randomized controlled trial in PPMS, to investigate the reliability or the ‘noise’ inherent in the EDSS, T25FW and nine-hole peg test (NHPT)<sup>6</sup> and their combinations.

Multiple Sclerosis Journal  
2021, Vol. 27(12) 1864–1874

DOI: 10.1177/  
1352458520987539

© The Author(s), 2021.



Article reuse guidelines:  
sagepub.com/journals-  
permissions

Correspondence to:

**MW Koch**

Department of Clinical  
Neurosciences, University of  
Calgary, Room 178, Heritage  
Medical Research Building,  
3330 Hospital Drive NW,  
Calgary, Alberta T2N 4N1,  
Canada.

[mwkocho@ucalgary.ca](mailto:mwkocho@ucalgary.ca)

**Marcus W Koch**

Department of Clinical  
Neurosciences, University  
of Calgary, Calgary, AB,  
Canada/Department of  
Community Health Sciences,  
University of Calgary,  
Calgary, AB, Canada

**Jop P Mostert**

Department of Neurology,  
Rijnstate Hospital, Arnhem,  
The Netherlands

**Bernard Uitdehaag**

Department of Neurology,  
Amsterdam University  
Medical Centers, Amsterdam,  
The Netherlands

**Gary Cutter**

Department of Biostatistics,  
University of Alabama at  
Birmingham, Birmingham,  
AL, USA

**Table 1.** Baseline characteristics.

|  | Placebo          | Fingolimod 0.5 mg | Fingolimod 1.25 mg | Total            |
|--|------------------|-------------------|--------------------|------------------|
| <i>n</i>   | 487              | 336               | 147                | 970              |
| Sex (F/M)  | 235/252          | 163/173           | 71/76              | 469/501          |
| Age (mean, SD)   | 49.0, 8.3        | 49.1, 8.6         | 48.3, 8.5          | 48.9, 8.5        |
| Disease duration since first symptoms (mean, SD)           | 5.9, 2.4         | 5.8, 2.5          | 5.9, 2.5           | 5.9, 2.4         |
| Disease duration since diagnosis (mean, SD)                | 2.9, 2.3         | 2.8, 2.6          | 2.7, 2.2           | 2.8, 2.4         |
| Patients with gadolinium-enhancing lesions ( <i>n</i> , %) | 64, 13.1%        | 46, 13.6%         | 17, 11.6%          | 127, 13.1%       |
| EDSS at baseline (median, IQR)                             | 4.5 (4.0–6.0)    | 4.5 (3.75–5.5)    | 4.5 (4.0–6.0)      | 4.5 (4.0–5.5)    |
| T25FW at baseline (median, IQR)                            | 6.9 (5.5–9.6)    | 7.1 (5.6–10.1)    | 7.0 (5.5–8.8)      | 7.0 (5.5–9.5)    |
| NHPT at baseline (Median, IQR)                             | 25.5 (22.5–31.1) | 25.9 (23.1–30.7)  | 25.6 (22.2–31.4)   | 25.7 (22.7–31.1) |
| PASAT at baseline (Median, IQR)                            | 48 (38–56)       | 49 (36–55)        | 50 (37–56)         | 49 (37–55)       |

SD: standard deviation; EDSS: expanded disability status scale; NHPT: nine-hole peg test; PASAT: paced auditory serial addition test; T25FW: timed 25-foot walk; IQR: interquartile range.

In addition to the increasing physical disability, PPMS is also characterized by progressive cognitive decline,<sup>7</sup> and it would be useful to have a reliable clinical outcome measure reflecting this aspect of the disease. The paced auditory serial addition test (PASAT) was previously the most widely used cognitive outcome and is still often used in cross-sectional studies. It is part of the multi-dimensional multiple sclerosis functional composite,<sup>8</sup> but its value as a longitudinal measure is debatable because of a large practice effect that can limit its use as a repeated measure.<sup>9</sup> Despite its popularity, we know relatively little about changes on the PASAT over time and about its usefulness as a trial outcome. In this study, we additionally investigated the value of the PASAT as a measure of disease progression in PPMS.

To aid the selection of appropriate eligibility criteria for PPMS trials, we also investigated the association of baseline factors with the risk of disability worsening over the course of 3 years of follow-up. Our analyses inform the selection of the most appropriate outcome measures and eligibility criteria for clinical trials in PPMS.

## Methods

### *Trial data set and ethics*

The INFORMS data set was obtained from Novartis (Novartis, Basel, Switzerland), the pharmaceutical company which conducted and oversaw the INFORMS

trial. The ethical approval for INFORMS is described in the original publication.<sup>5</sup> The University of Calgary Conjoint Health Research Ethics Board Ethical granted ethical approval for this analysis. INFORMS was a randomized, double-blind, placebo controlled trial conducted at 148 centres in 18 countries. Key inclusion criteria were age: 25–65 years, a clinical diagnosis of PPMS, disease progression for 1 year or more, a disease duration of 2–10 years and objective evidence of disability worsening in the 2 years before inclusion. In INFORMS, participants were initially randomly assigned to receive either fingolimod 1.25 mg per day or placebo, but during the trial, the decision was made to discontinue the development of fingolimod 1.25 mg and to continue with fingolimod 0.5 mg instead. Participants who had been assigned to 1.25 mg were switched to 0.5 mg providing for variable exposure to 1.25 mg. We present the group of patients originally assigned to fingolimod 1.25 mg as separate group in Table 1. For certain analyses, we combined both fingolimod groups into a single group. For the presentation of disease duration from onset and from the time of diagnosis, we imputed a missing day of the month as the 15th of the month, and a missing month as July of the year.

### *Progression rates*

We determined the proportion of individuals with disability worsening and improvement by comparing baseline and follow-up disability measures. Patients

missing the disability measure at baseline, the time point of interest or the corresponding confirmation assessment (at either 3 or 6 months subsequently) were excluded from these analyses. Disability worsening and improvement were defined as a 20% or more worsening/improvement from baseline in the time for the T25FW<sup>4</sup> and the NHPT.<sup>6</sup> According to the definition of historical trials in SPMS<sup>10–13</sup> and PPMS,<sup>1</sup> we defined worsening/improvement on the EDSS as an increase/decrease of one whole point on the EDSS if the baseline EDSS was 5.5 or lower, and of one half point if the baseline EDSS was 6.0 or 6.5. Since no agreed-upon definitions of significant worsening exist for the PASAT, we calculated mean PASAT-3 scores throughout follow-up. We also investigated worsening and improvement on the PASAT-3 from baseline (1) by any degree, (2) by at least four points and (3) by at least 20%.

#### *Investigations of ‘noise’*

We investigated the reliability or ‘noise’ inherent in the EDSS, T25FW and NHPT in three ways.

*Unconfirmed versus confirmed disability worsening.* First, we compared unconfirmed and ‘confirmed’ disability worsening. We labelled a worsening event ‘confirmed’ if (1) a disability measure showed significant worsening compared to baseline and (2) was confirmed as worsened at a confirmation measurement 3 or 6 months. An ideal robust clinical outcome in PPMS should have only a small difference between unconfirmed and confirmed disability worsening.

*Confirmed versus sustained disability worsening.* We also compared ‘confirmed’ with ‘sustained’ disability worsening. We labelled a worsening event ‘sustained’ if (1) a disability measure showed significant worsening compared to baseline, (2) remained significantly worsened at a confirmation measurement 3 or 6 months and (3) remained significantly worsened at the last (36 month) trial visit. An ideal outcome of irreversible disability worsening in PPMS should have only a small difference between confirmed and sustained disability worsening. However, one disadvantage of this approach is that the length of the ‘sustained’ period varies depending on when the index worsening first occurs.

*Disability worsening versus similarly defined improvement.* To investigate whether the measure truly is a reflection of the ongoing worsening of disability, we compared disability worsening with similarly defined improvement. An outcome measuring the chronically progressive disease process of PPMS

should have only a small proportion of patients ‘improving’ on an outcome measure, and a large proportion of patients worsening. The proportion of patients with disability worsening should increase over time, reflecting chronic progression in PPMS, while the proportion of patients with improvement should decline or remain unchanged over time.

#### *Baseline factors associated with disability worsening*

To aid in the selection of inclusion criteria for clinical trials, we investigated the association of baseline characteristics with disease progression at 12, 24 and 36 months of follow-up using logistic regression models. We used worsening on the EDSS or T25FW (unconfirmed and 3 month confirmed) at 12, 24 and 36 months as the dependent (outcome) variables and age, disease duration, sex, treatment (placebo or fingolimod), EDSS score at baseline, T25FW at baseline and contrast enhancing lesions on the screening MRI (present or absent) as the independent (predictor) variables. Statistical significance was taken to be at the two-tailed 0.05 level. All statistical analyses were performed with the R statistical software package for Windows, version 4.0.2.<sup>14</sup>

#### *Data availability*

The data used in this study are available upon request from Novartis. Individual participant data collected during the trial will be shared after anonymization and on approval of a research proposal and data sharing agreement. Research proposals can be submitted online (<https://www.clinicalstudydatarequest.com>).

## **Results**

#### *INFORMS data set*

The INFORMS data set contained individual patient level data of 970 participants. Table 1 shows their baseline characteristics.

#### *Progression rates*

Table 2 shows the proportion of trial participants with unconfirmed and confirmed disability worsening over the course of the trial. The T25FW had the highest number of worsening events over time, followed by the EDSS. The NHPT showed the lowest progression rates over follow-up. To explore whether the NHPT may be a more useful outcome in participants with advanced disability, we investigated NHPT worsening in a subgroup of patients with a baseline EDSS of

**Table 2.** Percentage of trial participants with unconfirmed and confirmed disability worsening throughout follow-up.

|                              | 3    | 6    | 9    | 12   | 15   | 18   | 21   | 24   | 27   | 30   | 33   | 36   |
|------------------------------|------|------|------|------|------|------|------|------|------|------|------|------|
| <b>EDSS</b>                  |      |      |      |      |      |      |      |      |      |      |      |      |
| UDP                          | 12.4 | 17.7 | 21.7 | 26.4 | 31.1 | 33.2 | 35.3 | 38.3 | 40.4 | 43.5 | 44.3 | 46.5 |
| <i>n</i>                     | 884  | 851  | 825  | 792  | 759  | 733  | 719  | 690  | 668  | 650  | 650  | 516  |
| 3M CDP                       | 8.8  | 12.4 | 17.0 | 21.7 | 26.5 | 28.0 | 29.8 | 33.2 | 35.7 | 39.3 | 36.5 |      |
| Percentage of UDP            | 71.0 | 70.1 | 78.3 | 82.2 | 85.2 | 84.3 | 84.4 | 86.7 | 88.4 | 90.3 | 82.4 | 0.0  |
| EDSS 6M CDP                  | 7.6  | 12.6 | 16.9 | 20.9 | 24.9 | 26.0 | 29.2 | 32.2 | 35.2 | 34.9 |      |      |
| Percentage of UDP            | 61.3 | 71.2 | 77.9 | 79.2 | 80.1 | 78.3 | 82.7 | 84.1 | 87.1 | 80.2 |      |      |
| <b>T25FW</b>                 |      |      |      |      |      |      |      |      |      |      |      |      |
| UDP                          | 16.2 | 22.7 | 28.6 | 33.0 | 37.2 | 41.3 | 44.4 | 45.5 | 47.4 | 47.2 | 48.3 | 52.1 |
| <i>n</i>                     | 846  | 823  | 796  | 757  | 717  | 683  | 664  | 642  | 620  | 602  | 590  | 466  |
| 3M CDP                       | 8.7  | 15.5 | 19.6 | 23.1 | 27.3 | 32.2 | 35.0 | 37.1 | 37.2 | 37.3 | 36.2 |      |
| Percentage of UDP            | 53.7 | 68.3 | 68.5 | 70.0 | 73.4 | 78.0 | 78.8 | 81.5 | 78.5 | 79.0 | 74.9 |      |
| 6M CDP                       | 8.8  | 14.7 | 18.5 | 22.5 | 27.6 | 32.0 | 33.9 | 35.4 | 38.1 | 34.0 |      |      |
| Percentage of UDP            | 54.3 | 64.8 | 64.7 | 68.2 | 74.2 | 77.5 | 76.4 | 77.8 | 80.4 | 72.0 |      |      |
| <b>NHPT</b>                  |      |      |      |      |      |      |      |      |      |      |      |      |
| UDP                          | 5.2  | 6.9  | 8.9  | 10.3 | 12.1 | 15.7 | 16.4 | 18.5 | 17.4 | 20.0 | 21.8 | 21.3 |
| <i>n</i>                     | 872  | 843  | 816  | 787  | 751  | 720  | 696  | 681  | 649  | 639  | 629  | 502  |
| 3M CDP                       | 1.8  | 3.3  | 4.5  | 5.8  | 7.5  | 8.8  | 11.0 | 11.2 | 11.7 | 14.8 | 12.1 |      |
| Percentage of UDP            | 34.6 | 47.8 | 50.6 | 56.3 | 62.0 | 56.1 | 67.1 | 60.5 | 67.2 | 74.0 | 55.5 |      |
| 6M CDP                       | 1.7  | 3.1  | 3.7  | 5.9  | 7.3  | 9.5  | 8.3  | 11.1 | 11.7 | 11.6 |      |      |
| Percentage of UDP            | 32.7 | 44.9 | 41.6 | 57.3 | 60.3 | 60.5 | 50.6 | 60.0 | 67.2 | 58.0 |      |      |
| <b>EDSS or T25FW</b>         |      |      |      |      |      |      |      |      |      |      |      |      |
| UDP                          | 25.7 | 32.6 | 39.4 | 45.0 | 49.4 | 52.7 | 55.6 | 57.4 | 59.4 | 61.2 | 61.8 | 65.9 |
| <i>n</i>                     | 847  | 831  | 810  | 774  | 745  | 719  | 703  | 681  | 657  | 639  | 641  | 508  |
| 3M CDP                       | 15.8 | 23.8 | 29.3 | 35.0 | 40.2 | 44.4 | 47.5 | 50.1 | 50.8 | 53.2 | 51.4 |      |
| Percentage of UDP            | 61.5 | 73.0 | 74.4 | 77.8 | 81.4 | 84.3 | 85.4 | 87.3 | 85.5 | 86.9 | 83.2 |      |
| 6M CDP                       | 15.2 | 23.4 | 28.5 | 34.8 | 39.8 | 43.2 | 46.2 | 48.4 | 51.2 | 49.2 |      |      |
| Percentage of UDP            | 59.1 | 71.8 | 72.3 | 77.3 | 80.6 | 82.0 | 83.1 | 84.3 | 86.2 | 80.4 |      |      |
| <b>EDSS or T25FW or NHPT</b> |      |      |      |      |      |      |      |      |      |      |      |      |
| UDP                          | 28.9 | 36.7 | 43.3 | 48.6 | 54.4 | 57.1 | 60.6 | 61.5 | 63.2 | 65.2 | 66.1 | 69.0 |
| <i>n</i>                     | 848  | 831  | 811  | 777  | 747  | 721  | 705  | 683  | 658  | 643  | 641  | 509  |
| 3M CDP                       | 17.3 | 26.1 | 31.4 | 38.1 | 43.7 | 47.9 | 51.6 | 53.3 | 54.1 | 57.1 | 55.6 |      |
| Percentage of UDP            | 59.9 | 71.1 | 72.5 | 78.4 | 80.3 | 83.9 | 85.1 | 86.7 | 85.6 | 87.6 | 84.1 |      |
| 6M CDP                       | 16.2 | 25.6 | 30.6 | 37.4 | 43.4 | 47.1 | 49.2 | 52.2 | 54.6 | 52.6 |      |      |
| Percentage of UDP            | 56.1 | 69.8 | 70.7 | 77.0 | 79.8 | 82.5 | 81.2 | 84.9 | 86.4 | 80.7 |      |      |

EDSS: expanded disability status scale; NHPT: nine-hole peg test; UDP: unconfirmed disability progression, CDP: confirmed disability progression, 3M: 3 months, 6M: 6 months; *n*: individuals with an available measurement; T25FW: timed 25-foot walk.

6.0 or higher ( $n=231$ ). This exploration showed only slightly increased progression rates compared to the whole cohort (e.g. 24-month unconfirmed disability progression (UDP) of 23.6% compared to 18.5% in the whole cohort and 24-month 3M CDP of 15.2% compared to 11.2% in the whole cohort; further data not shown). Table 6 and Figure 2 show mean PASAT scores and patients with worsening and improvement on the PASAT throughout follow-up. Mean PASAT scores steadily increased, especially within the first year, and then changed little until the end of the trial.

Throughout the trial, participants were more likely to improve on the PASAT than to worsen, using any of the three investigated definitions of worsening and improvement.

#### *Investigations of 'noise'*

*Unconfirmed versus confirmed disability worsening.* Table 2 shows the difference between unconfirmed and confirmed disability worsening on single and combined outcome measures. The EDSS showed

**Table 3.** Comparison of confirmed versus sustained disability worsening.

|                              | 3    | 6    | 9    | 12   | 15   | 18   | 21   | 24   | 27   | 30   |
|------------------------------|------|------|------|------|------|------|------|------|------|------|
| <b>EDSS</b>                  |      |      |      |      |      |      |      |      |      |      |
| 3M CDP                       | 8.8  | 12.4 | 17.0 | 21.7 | 26.5 | 28.0 | 29.8 | 33.2 | 35.7 | 39.3 |
| 3M SDP                       | 5.5  | 7.2  | 10.1 | 14.2 | 18.3 | 19.7 | 21.9 | 25.4 | 29.3 | 33.4 |
| Percentage of CDP            | 62.5 | 58.1 | 59.4 | 65.4 | 69.1 | 70.4 | 73.5 | 76.5 | 82.1 | 85.0 |
| <b>T25FW</b>                 |      |      |      |      |      |      |      |      |      |      |
| 3M CDP                       | 8.7  | 15.5 | 19.6 | 23.1 | 27.3 | 32.2 | 35.0 | 37.1 | 37.2 | 37.3 |
| 3M SDP                       | 2.7  | 6.1  | 8.1  | 11.3 | 17.2 | 20.2 | 23.1 | 26.9 | 28.4 | 29.6 |
| Percentage of CDP            | 31.0 | 39.4 | 41.3 | 48.9 | 63.0 | 62.7 | 66.0 | 72.5 | 76.3 | 79.4 |
| <b>NHPT</b>                  |      |      |      |      |      |      |      |      |      |      |
| 3M CDP                       | 1.8  | 3.3  | 4.5  | 5.8  | 7.5  | 8.8  | 11.0 | 11.2 | 11.7 | 14.8 |
| 3M SDP                       | 0.5  | 1.0  | 1.9  | 2.2  | 3.8  | 4.3  | 5.4  | 6.5  | 7.6  | 9.2  |
| Percentage of CDP            | 27.8 | 30.3 | 42.2 | 37.9 | 50.7 | 48.9 | 49.1 | 58.0 | 65.0 | 62.2 |
| <b>EDSS or T25FW</b>         |      |      |      |      |      |      |      |      |      |      |
| 3M CDP                       | 15.8 | 23.8 | 29.3 | 35.0 | 40.2 | 44.4 | 47.5 | 50.1 | 50.8 | 53.2 |
| 3M SDP                       | 8.4  | 12.4 | 16.7 | 22.3 | 28.3 | 32.0 | 36.2 | 40.4 | 42.3 | 45.4 |
| Percentage of CDP            | 53.2 | 52.1 | 57.0 | 63.7 | 70.4 | 72.1 | 76.2 | 80.6 | 83.3 | 85.3 |
| <b>EDSS or T25FW or NHPT</b> |      |      |      |      |      |      |      |      |      |      |
| 3M CDP                       | 17.3 | 26.1 | 31.4 | 38.1 | 43.7 | 47.9 | 51.6 | 53.3 | 54.1 | 57.1 |
| 3M SDP                       | 8.9  | 13.5 | 17.8 | 24.3 | 30.8 | 34.7 | 39.2 | 43.2 | 45.3 | 48.8 |
| Percentage of CDP            | 51.4 | 51.7 | 56.7 | 63.8 | 70.5 | 72.4 | 76.0 | 81.1 | 83.7 | 85.5 |

EDSS: expanded disability status scale; NHPT: nine-hole peg test; CDP: confirmed disability progression, SDP: sustained disability progression, 3M: 3 months, 6M: 6 months; T25FW: timed 25-foot walk.

the lowest difference between unconfirmed and confirmed disability worsening, with a large majority of unconfirmed worsening events confirmed at 3 or 6 months (e.g. 82.2% of those with unconfirmed 12-month worsening events were confirmed at 3 months). For the T25FW this difference was slightly larger (70% of unconfirmed 12-month worsening events confirmed at 3 months), and largest for the NHPT (with only 56.3% of unconfirmed 12-month worsening events confirmed at 3 months, Table 2). There was little difference between 3 and 6 month confirmation, although 6 month was slightly lower with regards to the difference between unconfirmed and confirmed disability worsening.

*Confirmed versus sustained disability worsening.* Table 3 shows the difference between confirmed and sustained worsening events for single and combined outcomes. The EDSS showed the lowest difference between confirmed and sustained disability worsening, with, for example, 65.4% of confirmed 12-month worsening events sustained until the end of the trial, followed by the T25FW (48.9%). The NHPT had the largest discrepancy between confirmed and sustained worsening events, with only 37.9% of confirmed 12-month worsening events sustained until the end of the trial (Table 3).

*Disability worsening versus similarly defined improvement.* Table 4 and Figure 1 show a comparison of worsening events with similarly defined improvement on single and combined outcome measures. Overall, improvement events were much rarer than worsening events, and remained stable throughout the course of the trial. The EDSS had the highest number of improvement events with around 10% of patients experiencing improvement (unconfirmed) on the EDSS, followed by the T25FW (with around 7% of improvement) and the NHPT (with around 4% of improvement; Table 4).

#### *Baseline factors associated with disability worsening*

EDSS and T25FW at baseline were consistently associated with EDSS and T25FW worsening at 12, 24 and 36 (or 33) months. Male sex and disease duration were associated with worsening on the EDSS in some but not all regression models (Table 5). Age, the presence of contrast enhancing lesions and fingolimod treatment were not associated with the risk of EDSS and T25FW disability worsening in any of the regression models. We performed these analyses with the treatment variable dichotomized into placebo or fingolimod, repeating analyses with separate 0.5 and 1.25 mg fingolimod arms did not change the results.



**Table 4.** Percentages of trial participants with disability worsening versus similarly defined improvement throughout follow-up.

|                              | 3    | 6    | 9    | 12   | 15   | 18   | 21   | 24   | 27   | 30   | 33   | 36   |
|------------------------------|------|------|------|------|------|------|------|------|------|------|------|------|
| <b>EDSS</b>                  |      |      |      |      |      |      |      |      |      |      |      |      |
| Worse UDP                    | 12.4 | 17.7 | 21.7 | 26.4 | 31.1 | 33.2 | 35.3 | 38.3 | 40.4 | 43.5 | 44.3 | 46.5 |
| Better UDP                   | 8.3  | 9.5  | 9.8  | 11.2 | 10.9 | 10.1 | 10.0 | 9.1  | 9.0  | 9.4  | 9.5  | 9.5  |
| Percentage of worse          | 66.9 | 53.7 | 45.2 | 42.4 | 35.0 | 30.4 | 28.3 | 23.8 | 22.3 | 21.6 | 21.4 | 20.4 |
| Worse 3M CDP                 | 8.8  | 12.4 | 17.0 | 21.7 | 26.5 | 28.0 | 29.8 | 33.2 | 35.7 | 39.3 | 36.5 |      |
| Better 3M CDP                | 3.8  | 4.9  | 5.8  | 7.5  | 6.9  | 6.7  | 6.2  | 6.4  | 6.6  | 6.7  | 5.3  |      |
| Percentage of worse          | 43.2 | 39.5 | 34.1 | 34.6 | 26.0 | 23.9 | 20.8 | 19.3 | 18.5 | 17.0 | 14.5 |      |
| <b>T25FW</b>                 |      |      |      |      |      |      |      |      |      |      |      |      |
| Worse UDP                    | 16.2 | 22.7 | 28.6 | 33.0 | 37.2 | 41.3 | 44.4 | 45.5 | 47.4 | 47.2 | 48.3 | 52.1 |
| Better UDP                   | 6.0  | 7.0  | 7.0  | 5.7  | 6.6  | 6.3  | 6.8  | 6.7  | 6.9  | 6.3  | 6.1  | 6.4  |
| Percentage of worse          | 37.0 | 30.8 | 24.5 | 17.3 | 17.7 | 15.3 | 15.3 | 14.7 | 14.6 | 13.3 | 12.6 | 12.3 |
| Worse 3M CDP                 | 8.7  | 15.5 | 19.6 | 23.1 | 27.3 | 32.2 | 35.0 | 37.1 | 37.2 | 37.3 | 36.2 |      |
| Better 3M CDP                | 2.5  | 3.8  | 3.7  | 3.5  | 3.8  | 4.5  | 4.8  | 4.1  | 4.1  | 4.7  | 3.9  |      |
| Percentage of worse          | 28.7 | 24.5 | 18.9 | 15.2 | 13.9 | 14.0 | 13.7 | 11.1 | 11.0 | 12.6 | 10.8 |      |
| <b>NHPT</b>                  |      |      |      |      |      |      |      |      |      |      |      |      |
| Worse UDP                    | 5.2  | 6.9  | 8.9  | 10.3 | 12.1 | 15.7 | 16.4 | 18.5 | 17.4 | 20.0 | 21.8 | 21.3 |
| Better UDP                   | 3.2  | 3.0  | 3.6  | 2.9  | 3.6  | 4.0  | 4.0  | 4.3  | 3.9  | 3.8  | 3.7  | 5.2  |
| Percentage of worse          | 61.5 | 43.5 | 40.4 | 28.2 | 29.8 | 25.5 | 24.4 | 23.2 | 22.4 | 19.0 | 17.0 | 24.4 |
| Worse 3M CDP                 | 1.8  | 3.3  | 4.5  | 5.8  | 7.5  | 8.8  | 11.0 | 11.2 | 11.7 | 14.8 | 12.1 |      |
| Better 3M CDP                | 1.5  | 1.4  | 1.7  | 2.0  | 1.8  | 2.0  | 2.1  | 2.4  | 2.2  | 1.5  | 2.5  |      |
| Percentage of worse          | 83.3 | 42.4 | 37.8 | 34.5 | 24.0 | 22.7 | 19.1 | 21.4 | 18.8 | 10.1 | 20.7 |      |
| <b>EDSS or T25FW</b>         |      |      |      |      |      |      |      |      |      |      |      |      |
| Worse UDP                    | 25.7 | 32.6 | 39.4 | 45.0 | 49.4 | 52.7 | 55.6 | 57.4 | 59.4 | 61.2 | 61.8 | 65.9 |
| Better UDP                   | 13.9 | 16.0 | 16.2 | 16.4 | 17.0 | 15.8 | 16.2 | 14.6 | 15.4 | 15.2 | 15.0 | 15.8 |
| Percentage of worse          | 54.1 | 49.1 | 41.1 | 36.4 | 34.4 | 30.0 | 29.1 | 25.4 | 25.9 | 24.8 | 24.3 | 24.0 |
| Worse 3M CDP                 | 15.8 | 23.8 | 29.3 | 35.0 | 40.2 | 44.4 | 47.5 | 50.1 | 50.8 | 53.2 | 51.4 |      |
| Better 3M CDP                | 6.4  | 8.4  | 9.2  | 10.8 | 10.4 | 10.8 | 10.5 | 10.0 | 10.3 | 10.8 | 9.5  |      |
| Percentage of worse          | 40.5 | 35.3 | 31.4 | 30.9 | 25.9 | 24.3 | 22.1 | 20.0 | 20.3 | 20.3 | 18.5 |      |
| <b>EDSS or T25FW or NHPT</b> |      |      |      |      |      |      |      |      |      |      |      |      |
| Worse UDP                    | 28.9 | 36.7 | 43.3 | 48.6 | 54.4 | 57.1 | 60.6 | 61.5 | 63.2 | 65.2 | 66.1 | 69.0 |
| Better UDP                   | 16.5 | 18.1 | 19.1 | 18.7 | 19.9 | 19.2 | 19.3 | 18.4 | 18.6 | 18.2 | 17.8 | 19.9 |
| Percentage of worse          | 57.1 | 49.3 | 44.1 | 38.5 | 36.6 | 33.6 | 31.8 | 29.9 | 29.4 | 27.9 | 26.9 |      |
| Worse 3M CDP                 | 17.3 | 26.1 | 31.4 | 38.1 | 43.7 | 47.9 | 51.6 | 53.3 | 54.1 | 57.1 | 55.6 |      |
| Better 3M CDP                | 7.7  | 9.6  | 10.6 | 12.5 | 11.9 | 12.7 | 12.5 | 12.0 | 11.9 | 11.8 | 11.4 |      |
| Percentage of worse          | 44.5 | 36.8 | 33.8 | 32.8 | 27.2 | 26.5 | 24.2 | 22.5 | 22.0 | 20.7 | 20.5 |      |

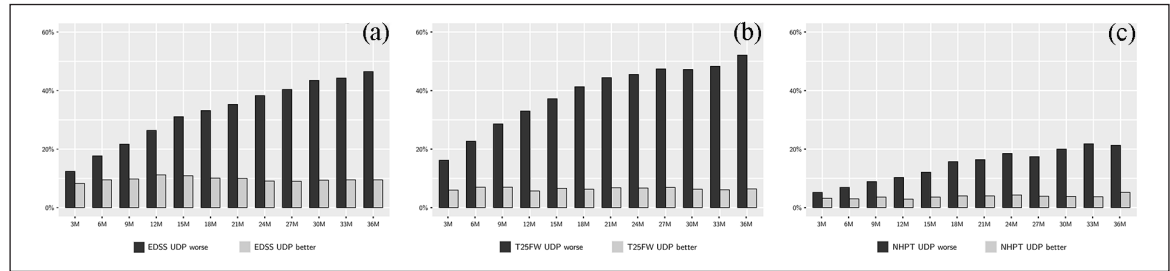
EDSS: expanded disability status scale; NHPT: nine-hole peg test; UDP: unconfirmed disability progression, CDP: confirmed disability progression, 3M: 3 months, 6M: 6 months; T25FW: timed 25-foot walk.

## Discussion

An ideal clinical outcome measure in PPMS should show a steadily growing number of worsening events over time in a disease that has no disease modifying treatments. These worsening events should reflect irreversible disability, so that the difference between raw and confirmed worsening, on one hand, and confirmed and sustained worsening, on the other hand, should be as low as possible. So far, there has been little motivation to compare outcome measures in

PPMS and other disease courses, because the agreed-upon standard for outcome measurement in all forms of MS has been the EDSS.

Our investigation of progression rates shows that—similar to our previous investigations in SPMS<sup>15</sup> and PPMS,<sup>16</sup> the T25FW had the highest proportion of patients with disability worsening over time, followed by the EDSS. The NHPT showed the lowest worsening rates, with only 12.1% of patients experiencing (3 month



**Figure 1.** Proportion of patients with disability worsening versus similarly defined (unconfirmed) improvement on the (a) EDSS, (b) T25FW and (c) NHPT.

**Table 5.** Results of the logistic regression models.

| Dependent variable           | Significant independent variables | Odds ratio (95% confidence interval) | <i>p</i> | AIC    |
|------------------------------|-----------------------------------|--------------------------------------|----------|--------|
| T25FW worse UDP 12 months    | EDSS at baseline                  | 1.44 (1.20–1.74)                     | 0.0001   | 944.79 |
| T25FW worse UDP 24 months    | EDSS at baseline                  | 1.57 (1.29–1.92)                     | <0.0001  | 867.77 |
| T25FW worse UDP 36 months    | Disease duration                  | 0.92 (0.85–0.99)                     | 0.04     | 631.07 |
|                              | T25FW at baseline                 | 0.93 (0.89–0.98)                     | 0.006    |        |
|                              | EDSS at baseline                  | 1.68 (1.32–2.14)                     | <0.0001  |        |
| T25FW worse 3M CDP 12 months | EDSS at baseline                  | 1.52 (1.24–1.88)                     | <0.0001  | 780.8  |
| T25FW worse 3M CDP 24 months | EDSS at baseline                  | 1.73 (1.40–2.14)                     | <0.0001  | 793.81 |
| T25FW worse 3M CDP 33 months | T25FW at baseline                 | 0.94 (0.89–0.99)                     | 0.03     | 683.76 |
|                              | EDSS at baseline                  | 1.50 (1.19–1.90)                     | 0.0006   |        |
| EDSS worse UDP 12 months     | Disease duration                  | 0.93 (0.87–0.99)                     | 0.04     | 862.32 |
|                              | Male sex                          | 1.63 (1.17–2.30)                     | 0.004    |        |
|                              | T25FW at baseline                 | 1.07 (1.04–1.12)                     | <0.0001  |        |
| EDSS worse UDP 24 months     | Male sex                          | 1.4 (1.01–1.93)                      | 0.04     | 877.6  |
|                              | T25FW at baseline                 | 1.04 (1.01–1.08)                     | 0.03     |        |
|                              | EDSS at baseline                  | 1.3 (1.07–1.57)                      | 0.007    |        |
| EDSS worse UDP 36 months     | T25FW at baseline                 | 1.06 (1.02–1.12)                     | 0.01     | 680.63 |
|                              | EDSS at baseline                  | 1.31 (1.05–1.64)                     | 0.02     |        |
| EDSS worse 3M CDP 12 months  | Male sex                          | 1.56 (1.09–2.25)                     | 0.02     | 774.53 |
|                              | T25FW at baseline                 | 1.08 (1.04–1.12)                     | <0.0001  |        |
| EDSS worse 3M CDP 24 months  | T25FW at baseline                 | 1.04 (1.01–1.08)                     | 0.03     | 821.77 |
|                              | EDSS at baseline                  | 1.31 (1.07–1.60)                     | 0.008    |        |
| EDSS worse 3M CDP 33 months  | T25FW at baseline                 | 1.06 (1.02–1.11)                     | 0.004    | 749.97 |
|                              | EDSS at baseline                  | 1.26 (1.02–1.56)                     | 0.03     |        |

EDSS: expanded disability status scale; AIC: Akaike information criterion; UDP: unconfirmed disability progression, CDP: confirmed disability progression; T25FW: timed 25-foot walk.

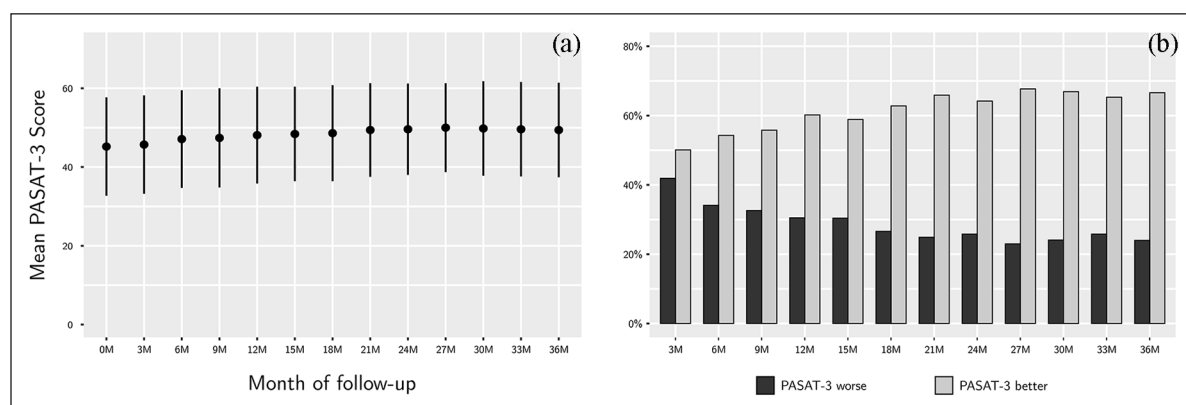
The models included sex, age, disease duration, EDSS at baseline, T25FW at baseline, contrast enhancing lesions on the screening MRI (present or absent) and the treatment arm (placebo or fingolimod) as independent (predictor) variables.

confirmed) disability worsening at 33 months. Limiting our investigation to a subgroup of patients with significant baseline disability (EDSS of 6.0 or greater, *n*=231), did not change this conclusion substantially. The NHPT is likely not a useful primary single outcome measure in PPMS, and we found no support for the idea that the NHPT would be more useful in patients with more advanced disability. Our investigation of the PASAT showed only little change in PASAT scores over the

course of this 3-year trial, except for a small increase in scores in especially the first year of the trial, which probably reflects a practice effect (Table 6 and Figure 2). Throughout the trial, participants on average were more likely to improve than to worsen on the PASAT, no matter which definition of worsening we chose, which is not in keeping with the chronically progressive cognitive decline that people with PPMS often experience. While this could in part be due to the worsening

**Table 6.** Change in the PASAT throughout the trial.

|   | 0    | 3    | 6    | 9    | 12   | 15   | 18   | 21   | 24   | 27   | 30   | 33   | 36   |
|---|------|------|------|------|------|------|------|------|------|------|------|------|------|
| PASAT-3   |      |      |      |      |      |      |      |      |      |      |      |      |      |
| Mean PASAT score  | 45.2 | 45.7 | 47.1 | 47.4 | 48.1 | 48.4 | 48.6 | 49.4 | 49.6 | 50.0 | 49.8 | 49.6 | 49.4 |
| SD PASAT score  | 12.5 | 12.5 | 12.4 | 12.6 | 12.3 | 12.0 | 12.2 | 11.9 | 11.6 | 11.3 | 12.0 | 12.0 | 12.0 |
| Participants with any worsening (%)                                 | –    | 41.9 | 34.1 | 32.6 | 30.5 | 30.4 | 26.6 | 24.9 | 25.8 | 23.0 | 24.1 | 25.8 | 24.0 |
| Participants with any improvement (%)                               | –    | 50.1 | 54.3 | 55.8 | 60.2 | 58.9 | 62.8 | 65.9 | 64.2 | 67.7 | 66.9 | 65.3 | 66.6 |
| Participants with four or more points worsening (%)                 | –    | 19.7 | 16.6 | 15.9 | 15.2 | 14.3 | 14.1 | 12.3 | 12.9 | 10.9 | 11.6 | 12.5 | 11.8 |
| Participants with four or more points improvement (%)               | –    | 24.8 | 32.6 | 34.0 | 35.7 | 35.8 | 42.1 | 41.4 | 43.8 | 43.5 | 45.4 | 43.2 | 46.6 |
| Participants with 20% or more worsening (%)                         | –    | 7.0  | 4.6  | 5.7  | 6.0  | 5.7  | 5.6  | 5.4  | 5.4  | 4.7  | 5.3  | 6.2  | 5.1  |
| Participants with 20% or more improvement (%)                       | –    | 10.1 | 14.7 | 14.1 | 16.3 | 16.8 | 21.4 | 20.5 | 21.6 | 20.6 | 21.8 | 21.8 | 22.8 |
| <i>n</i>  | 927  | 860  | 832  | 801  | 771  | 734  | 710  | 684  | 667  | 635  | 623  | 616  | 491  |
| PASAT: paced auditory serial addition test; SD: standard deviation. |      |      |      |      |      |      |      |      |      |      |      |      |      |



**Figure 2.** (a) Mean PASAT scores (error bars represent the standard deviation) throughout follow-up. The PASAT does not show worsening over time, but a slight increase in mean scores up to about 12 months, and little change afterwards. This slight increase in PASAT scores may be due to a practice effect. (b) Improvement of the PASAT compared to baseline is more likely than worsening throughout the trial.

patients dropping out of the trial, we did not see strong evidence for this. Given these results, the PASAT is likely not a useful primary or combined outcome measure in PPMS trials. The symbol digit modalities test (SDMT)<sup>17</sup> is now most often used as the primary cognitive outcome in MS, but was unfortunately not used in INFORMS. The value of SDMT as an outcome in PPMS should be investigated in other clinical trial data sets and clinical cohorts.

The primary goal of treatment in PPMS is the prevention or delay of irreversible disability. An outcome measure that reflects this irreversibility should therefore only show a small difference between unconfirmed and confirmed or sustained disability worsening. In the INFORMS data set, the EDSS showed the highest consistency between unconfirmed and confirmed and between confirmed and sustained

disability worsening, followed by the T25FW. Around two-thirds of 12-month (3 month confirmed) worsening events on the EDSS were sustained until 3 years of follow-up. This is in contrast to a seminal study in relapsing–remitting MS (RRMS), where only about half of all (3 months) confirmed worsening events at 1 year were sustained until 2 years of follow-up.<sup>18</sup> This difference between the earlier study in RRMS and our study may be because the RRMS study included participants with EDSS scores in the lower portion of the scale, where the EDSS is known to have poorer test–retest reliability.<sup>19,20</sup>

Our investigation of worsening versus similarly defined improvement is based on the idea that (in a trial that has not demonstrated a treatment effect) a useful outcome measure in PPMS should reflect the ongoing clinical worsening. While plateaus and



occasional improvements are possible in PPMS, the clinical picture is dominated by a slow and steady decline across all functional systems. Based on this reasoning, an ideal outcome measure in PPMS should show steady worsening of disability, while improvement by the same margin on the same outcome would then either be due to measurement error or a very rare 'true' improvement event; either way, worsening events should over time vastly outnumber improvement events in PPMS. Ebers et al.<sup>21</sup> showed that the EDSS improved about as often as it worsened in RRMS trial cohorts followed for shorter periods of time. Fortunately, this effect was much less noticeable in the INFORMS data set, with improvement rates of largely below 10% for the EDSS and even lower proportions for the other single outcome measures. Combining the EDSS and T25FW resulted in a larger proportion of individuals with disability worsening, without substantially increasing the differences between unconfirmed and confirmed or between confirmed and sustained disability worsening. Including the NHPT in a combined outcome added little to progression rates.

The EDSS is currently the standard primary outcome measure in all forms of MS. Our findings suggest that the T25FW would also be a good choice for primary outcome measure in PPMS. It may seem limiting to use a pure ambulation measure as the primary outcome, but it should be kept in mind that the EDSS is almost exclusively an ambulation measure at values of 4.0 and higher, which applies to the vast majority of the participants in INFORMS. The T25FW could be used in isolation or in combination with the EDSS. However, we base this recommendation on this investigation of a single trial data set. While INFORMS was a well conducted representative trial, the precise impact of using the T25FW as primary outcome measure on statistical power, sample size calculations, and trial duration should also be investigated in other PPMS trial data sets.

One reason for the relative high reliability of worsening events in INFORMS may lie in the lower inflammatory disease activity in PPMS compared to RRMS and SPMS. INFORMS included participants with notably low markers of inflammatory disease activity; for instance, only 13% of participants had contrast enhancing lesions at baseline, which is about half of, for example, the ORATORIO trial of ocrelizumab in PPMS (with 27% of patients having contrast enhancing lesions at baseline)<sup>1</sup> or comparable trials in SPMS: ASCEND (22%)<sup>22</sup> and IMPACT (34%).<sup>23</sup> This suggests that in the INFORMS cohort disability worsening is likely not driven by overt focal inflammatory disease activity.

We performed regression analyses on the association of baseline characteristics and the risk of disability worsening. Our models showed that age, disease duration, sex and having contrast enhancing lesions at baseline were not consistently associated with the risk of disability worsening in PPMS as compared to the frequent association in RRMS. This may be due to homogeneity of these risk factors within the study population or suggest that it may not be necessary to adjust for these factors by formulating specific eligibility criteria. In contrast to these findings, it appeared that younger patients with contrast enhancing lesions at baseline were more likely to benefit from immunomodulatory treatment in the ORATORIO trial, the only positive phase III trial in PPMS to date.<sup>1</sup> Through its eligibility criteria, INFORMS may have selected for a group of patients with meaningfully less focal inflammatory disease activity: the INFORMS cohort is, on average, 5 years older than the ORATORIO cohort, and included participants aged between 25 and 65 years, while ORATORIO included participants aged between 18 and 55 years. Focal inflammatory disease activity is in part an inverse function of age: two studies, which mostly included patients with RRMS and SPMS, showed that the proportion of patients with contrast enhancing lesions declines almost linearly as a function of age.<sup>24,25</sup> It would be worthwhile to explore the association of MRI characteristics, age and disability worsening in other natural history and clinical trial cohorts in PPMS.

### Declaration of Conflicting Interest

The author(s) declared the following potential conflicts of interest with respect to the research, authorship and/or publication of this article: M.W.K. received consulting fees and travel support from Biogen, Novartis, Roche, Sanofi Genzyme and EMD Serono. J.P.M. reports no disclosures. B.U. received consultancy fees and/or research support from Biogen, Sanofi Genzyme, EMD Serono, Novartis, Roche and Teva. G.C. served on Data and Safety Monitoring Boards: Astra-Zeneca, Avexis Pharmaceuticals, Biolinex, Brainstorm Cell Therapeutics, Bristol Meyers Squibb/Celgene, CSL Behring, Galmed Pharmaceuticals, Horizon Pharmaceuticals, Hisun Pharmaceuticals, Mapi Pharmaceuticals LTD, Merck, Merck/Pfizer, Opko Biologics, OncoImmune, Neurim, Novartis, Ophazyme, Sanofi-Aventis, Reata Pharmaceuticals, Teva Pharmaceuticals, VielaBio, Inc., Vivus, NHLBI (Protocol Review Committee) and NICHD (OPRU oversight committee). Consulting or Advisory Boards: Biodelivery Sciences International, Biogen, ClickTherapeutics, Genzyme, Genentech, GWPharmaceuticals, Immunic, Klein-Buendel Incorporated,

Medimmune, Medday, Neurogenesis LTD, Novartis, Osmotica Pharmaceuticals, Perception Neurosciences, Recursion/Cerexis Pharmaceuticals, Roche and TG Therapeutics. G.C. is employed by the University of Alabama at Birmingham and President of Pythagoras, Inc., a private consulting company located in Birmingham, AL.

### Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

### ORCID iD

Marcus W Koch  <https://orcid.org/0000-0001-9972-5092>

### References

- Montalban X, Hauser SL, Kappos L, et al. Ocrelizumab versus placebo in primary progressive multiple sclerosis. *N Engl J Med* 2017; 376(3): 209–220.
- Kurtzke JF. Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). *Neurology* 1983; 33(11): 1444–1452.
- Koch MW, Cutter G, Stys PK, et al. Treatment trials in progressive MS – Current challenges and future directions. *Nat Rev Neurol* 2013; 9(9): 496–503.
- Motl RW, Cohen JA, Benedict R, et al. Validity of the timed 25-foot walk as an ambulatory performance outcome measure for multiple sclerosis. *Mult Scler* 2017; 23(5): 704–710.
- Lublin F, Miller DH, Freedman MS, et al. Oral fingolimod in primary progressive multiple sclerosis (INFORMS): A phase 3, randomised, double-blind, placebo-controlled trial. *Lancet Lond Engl* 2016; 387(10023): 1075–1084.
- Feys P, Lamers I, Francis G, et al. The nine-hole peg test as a manual dexterity performance measure for multiple sclerosis. *Mult Scler* 2017; 23(5): 711–720.
- Chiaravallotti ND and DeLuca J. Cognitive impairment in multiple sclerosis. *Lancet Neurol* 2008; 7(12): 1139–1151.
- Cutter GR, Baier ML, Rudick RA, et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. *Brain* 1999; 122(Pt5): 871–882.
- Sumowski JF, Benedict R, Enzinger C, et al. Cognition in multiple sclerosis: State of the field and priorities for the future. *Neurology* 2018; 90(6): 278–288.
- Anon. Placebo-controlled multicentre randomised trial of interferon beta-1b in treatment of secondary progressive multiple sclerosis. European Study Group on Interferon Beta-1b in Secondary Progressive MS. *Lancet* 1998; 352(9139): 1491–1497.
- Hommes OR, Sørensen PS, Fazekas F, et al. Intravenous immunoglobulin in secondary progressive multiple sclerosis: Randomised placebo-controlled trial. *Lancet* 2004; 364(9440): 1149–1156.
- Panitch H, Miller A, Paty D, et al. Interferon beta-1b in secondary progressive MS: Results from a 3-year controlled study. *Neurology* 2004; 63(10): 1788–1795.
- Anon. Randomized controlled trial of interferon-beta-1a in secondary progressive MS: Clinical results. *Neurology* 2001; 56(11): 1496–1504.
- R Development Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2020, <http://www.R-project.org/>.
- Koch MW, Mostert J, Uitdehaag B, et al. Clinical outcome measures in SPMS trials: An analysis of the IMPACT and ASCEND original trial data sets. *Mult Scler Houndmills Basingstoke Engl* 2020; 26: 1540–1549.
- Koch MW, Cutter GR, Giovannoni G, et al. Comparative utility of disability progression measures in PPMS analysis of the PROMiSe data set. *Neurol Neuroimmunol Neuroinflamm* 2017; 4(4): e358.
- Benedict RH, DeLuca J, Phillips G, et al. Validity of the symbol digit modalities test as a cognition performance outcome measure for multiple sclerosis. *Mult Scler* 2017; 23(5): 721–733.
- Liu C and Blumhardt LD. Disability outcome measures in therapeutic trials of relapsing-remitting multiple sclerosis: Effects of heterogeneity of disease course in placebo cohorts. *J Neurol Neurosurg Psychiatry* 2000; 68(4): 450–457.
- Hobart J, Freeman J and Thompson A. Kurtzke scales revisited: The application of psychometric methods to clinical intuition. *Brain* 2000; 123(Pt5): 1027–1040.
- Goodkin DE, Cookfair D, Wende K, et al. Inter- and intrarater scoring agreement using grades 1.0 to 3.5 of the Kurtzke expanded disability status scale (EDSS). *Neurology* 1992; 42(4): 859–863.

21. Ebers GC, Heigenhauser L, Daumer M, et al. Disability as an outcome in MS clinical trials. *Neurology* 2008; 71(9): 624–631.
22. Kapoor R, Ho P-R, 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(5): 405–415.
23. Cohen JA, Cutter GR, Fischer JS, et al. Benefit of interferon beta-1a on MSFC progression in secondary progressive MS. *Neurology* 2002; 59(5): 679–687.
24. Tortorella C, Bellacosa A, Paolicelli D, et al. Age-related gadolinium-enhancement of MRI brain lesions in multiple sclerosis. *J Neurol Sci* 2005; 239(1): 95–99.
25. Koch MW, Mostert J, Greenfield J, et al. Gadolinium enhancement on cranial MRI in multiple sclerosis is age dependent. *J Neurol* 2020; 267(9): 2619–2624.

Visit SAGE journals online  
[journals.sagepub.com/  
home/msj](http://journals.sagepub.com/home/msj)

 SAGE journals