

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Kappos, L;Butzkueven, H;Wiendl, H;Spelman, T;Pellegrini, F;Chen, Y;Dong, Q;Koendgen, H;Belachew, S;Trojano, M

Title:

Greater sensitivity to multiple sclerosis disability worsening and progression events using a roving versus a fixed reference value in a prospective cohort study

Date:

2018-06-01

Citation:

Kappos, L., Butzkueven, H., Wiendl, H., Spelman, T., Pellegrini, F., Chen, Y., Dong, Q., Koendgen, H., Belachew, S. & Trojano, M. (2018). Greater sensitivity to multiple sclerosis disability worsening and progression events using a roving versus a fixed reference value in a prospective cohort study. *Multiple Sclerosis Journal*, 24 (7), pp.963-973. <https://doi.org/10.1177/1352458517709619>.

Persistent Link:

<https://hdl.handle.net/11343/259485>

License:

CC BY-NC

# Greater sensitivity to multiple sclerosis disability worsening and progression events using a roving versus a fixed reference value in a prospective cohort study

Ludwig Kappos, Helmut Butzkueven, Heinz Wiendl, Timothy Spelman, Fabio Pellegrini, Yi Chen, Qunming Dong, Harold Koendgen, Shibeshih Belachew and Maria Trojano; On Behalf of the Tysabri® Observational Program (TOP) Investigators

## Abstract

**Background:** Confirmed Expanded Disability Status Scale (EDSS) progression occurring after a fixed-study entry baseline is a common measure of disability increase in relapsing-remitting multiple sclerosis (RRMS) studies but may not detect all disability progression events, especially those unrelated to overt relapses.

**Objective:** To evaluate possible measures of disability progression unrelated to relapse using EDSS data over ~5.5 years from the Tysabri® Observational Program (TOP).

**Methods:** TOP is an ongoing, prospective, open-label study in RRMS patients receiving intravenous 300 mg natalizumab every 4 weeks. Measures of increasing disability were assessed using as a reference either study baseline score or a “roving” system that resets the reference score after  $\geq 24$ - or  $\geq 48$ -week confirmation of a new score.

**Results:** This analysis included 5562 patients. Approximately 70% more EDSS progression events unrelated to relapse and 50% more EDSS worsening events overall were detected with a roving reference score (cumulative probability: 17.6% and 29.7%, respectively) than with a fixed reference baseline score (cumulative probability: 10.1% and 20.3%, respectively).

**Conclusion:** In this long-term observational RRMS dataset, a roving EDSS reference value was more efficient than a study baseline EDSS reference in detecting progression/worsening events unrelated to relapses and thus the transition to secondary progressive disease.

**Keywords:** Multiple sclerosis, relapsing-remitting, natalizumab, disease progression, Expanded Disability Status Scale, secondary progressive multiple sclerosis

Date received: 21 December 2016; revised: 20 April 2017; accepted: 23 April 2017

## Introduction

Prevention of disability accumulation in patients with relapsing-remitting multiple sclerosis (RRMS) is a common goal in clinical trials and clinical practice.<sup>1</sup> Disability worsening is typically measured by increases in Expanded Disability Status Scale (EDSS) score confirmed at 12 or 24 weeks or time points later in the trial, using the baseline EDSS score as reference. Confirmation of EDSS increase at 12 or 24 weeks or a later time point reduces the likelihood of capturing events that may subsequently revert,<sup>2,3</sup> and the European Medicines Agency<sup>4</sup> has recently

recommended that disability worsening be confirmed by measurements taken at least 24 weeks apart.

Most RRMS studies enroll patients with clinical disease activity, and though the enrollment criteria of these studies require that the last relapse not have occurred within a certain timeframe before baseline (usually at least 30 days), a substantial fraction of patients will experience EDSS score regression,<sup>5</sup> especially during the first year, which may be associated with prolonged recovery from relapse. Such an initial decrease in EDSS score after treatment

Multiple Sclerosis Journal

2018, Vol. 24(7) 963–973

DOI: 10.1177/  
1352458517709619

© The Author(s), 2017.



Reprints and permissions:  
[http://www.sagepub.co.uk/  
journalsPermissions.nav](http://www.sagepub.co.uk/journalsPermissions.nav)

Correspondence to:

**L. Kappos**  
Neurologic Clinic and  
Polyclinic, Departments  
of Medicine, Clinical  
Research, Biomedicine and  
Biomedical Engineering,  
University Hospital of Basel,  
Petersgraben 4, CH-4031  
Basel, Switzerland.  
[ludwig.kappos@usb.ch](mailto:ludwig.kappos@usb.ch)

**Ludwig Kappos**  
Neurologic Clinic and  
Polyclinic, Departments of  
Medicine, Clinical Research,  
Biomedicine and Biomedical  
Engineering, University  
Hospital of Basel, Basel,  
Switzerland

**Helmut Butzkueven**  
Department of Medicine,  
Royal Melbourne Hospital,  
University of Melbourne,  
Parkville, VIC, Australia/  
Department of Neurology,  
Box Hill Hospital, Monash  
University, Box Hill, VIC,  
Australia

**Heinz Wiendl**  
Department of Neurology—  
Inflammatory Disorders of  
the Nervous System and  
Neurooncology, University  
of Münster, Münster,  
Germany

**Timothy Spelman**  
Department of Medicine,  
Royal Melbourne Hospital,  
University of Melbourne,  
Parkville, VIC, Australia

**Fabio Pellegrini**  
**Harold Koendgen**  
Biogen International GmbH,  
Zug, Switzerland

**Qunming Dong**  
**Yi Chen**  
**Shibeshih Belachew**  
Biogen, Cambridge, MA,  
USA

**Maria Trojano**  
Department of Basic Medical  
Sciences, Neuroscience and  
Sense Organs, University of  
Bari Aldo Moro, Bari, Italy

initiation may reduce the detection rate of subsequent events of disability worsening or progression, as patients first have to progress back to the baseline level and then beyond it to register a worsening or progression event.

The typical disease course of multiple sclerosis (MS) begins with clinically active RRMS, characterized by the occurrence of relapses, acute or subacute episodes of new or increasing neurologic dysfunction, followed by full or partial recovery in the absence of fever or infection.<sup>6</sup> In most RRMS patients, the disease eventually advances to a secondary progressive stage.<sup>7</sup> Secondary progressive MS (SPMS) is characterized by an initial relapsing-remitting disease course followed by progression of variable rate with or without occasional relapses, minor remissions, and plateaus.<sup>8</sup> The median time to secondary progression has been estimated as 15–19 years from RRMS onset.<sup>7,9</sup> Reaching SPMS appears to be the strongest determinant of poor long-term disease prognosis and is more dependent on age than on disease duration.<sup>10–14</sup>

Clear metrics for sensitive and reliable identification of the transition from RRMS to SPMS have been lacking.<sup>6</sup> A comprehensive analysis using the large MSBase cohort to evaluate potential definitions found the highest specificity in a definition requiring an EDSS score increase of  $\geq 1.0$  (or  $\geq 0.5$  from a baseline score  $\geq 6.0$ ), resulting in a minimum score of 4.0 in the absence of relapses that was confirmed after  $\geq 3$  months within the leading functional system, along with a minimum pyramidal functional system score of 2.0.<sup>15</sup> Despite its high specificity, this definition would only capture progressive disease in a rather advanced stage that may be less amenable to treatment.

When referring to increases in disability, recent literature has suggested that the term “worsening” be used in place of “progression” to describe increasing disability in patients with relapsing forms of the disease, with the term “progression” reserved for patients in the progressive phase of MS, defined by progressively increasing disability unrelated to relapse activity.<sup>6,15</sup>

In this study, we explore the use of a roving EDSS reference value to enhance detection of EDSS worsening events. The use of a roving EDSS reference value should also allow more sensitive measurement (in the total study population) of disability progression within relapse-free epochs according to specific, time-based interval definitions (e.g. 24 or 48 weeks

apart). Using data from a period of approximately 5.5 years in the Tysabri® Observational Program (TOP)<sup>16</sup> study of natalizumab-treated RRMS patients, we evaluate metrics using fixed and roving EDSS baseline criteria for identification of changes in disability.

## Materials and methods

### Study design

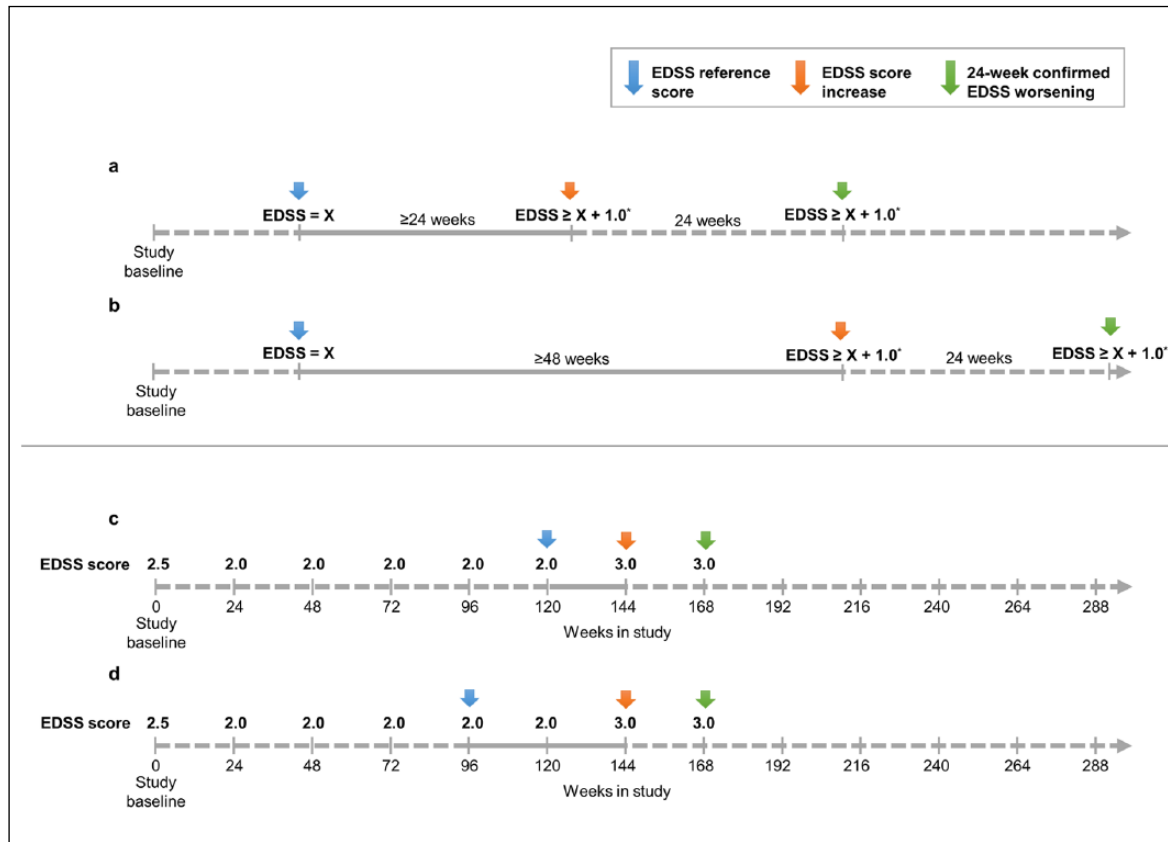
TOP (ClinicalTrials.gov: NCT00493298) is an ongoing, prospective, observational, 10-year open-label study of patients with RRMS in clinical practice settings in Europe, Australia, Argentina, and Canada.<sup>16</sup> The study protocol was approved by each center’s independent ethics committee. A complete list of investigators and the countries in which they practice is included in the Supplementary Material. The study design was written in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines, and all enrolled patients provided written informed consent.

The TOP methodology and interim conventional disability progression outcomes have been published.<sup>16</sup> Briefly, patients who have received  $\leq 3$  infusions of natalizumab prior to enrollment are eligible to enroll in TOP. Patients in TOP receive intravenous infusions of 300 mg natalizumab every 4 weeks.

### Assessments

EDSS scores were assessed at regular clinical visits approximately every 24 weeks. For the sake of clarity and consistency with the revised definitions of the clinical course of MS in Lublin et al.,<sup>6</sup> throughout this article, we refer to a disability increase associated with relapses as EDSS *worsening*, whereas EDSS *progression* is reserved for a disability increase unrelated to overt relapse activity. EDSS worsening and progression events were defined as increases in EDSS score of  $\geq 1.5$  points from an EDSS score of 0.0,  $\geq 1.0$  point from an EDSS score of 1.0–5.5, or  $\geq 0.5$  point from an EDSS score  $\geq 6.0$ . EDSS worsening and progression events were assessed using as a reference either the conventional fixed study baseline EDSS score or a roving EDSS score in which the increase or decrease had to be separated from the last EDSS assessment by at least 24 or 48 weeks (Figure 1). All EDSS worsening and progression events were required to be confirmed at 24 weeks.

Using the fixed space baseline EDSS score reference or the roving EDSS reference value, we



**Figure 1.** Schematic of the roving Expanded Disability Status Scale (EDSS) reference system. Confirmed EDSS worsening (a)  $\geq 24$  or (b)  $\geq 48$  weeks apart using a roving reference is illustrated. Also shown are hypothetical examples of confirmed EDSS worsening (c)  $\geq 24$  or (d)  $\geq 48$  weeks apart that would be captured using a roving EDSS reference and not accounted for using the conventional study baseline as EDSS.

\*Increase in EDSS score of  $\geq 1.5$  points from a score of 0.0,  $\geq 1.0$  point from a score of 1.0–5.5, or  $\geq 0.5$  points from a score  $\geq 6.0$ .

assessed both overall confirmed EDSS worsening and EDSS progression unrelated to relapse. An observed EDSS progression event was considered unrelated to any relapse if no concurrent relapse had been recorded from the 30 days prior to the reference EDSS assessment to either 30 days or 12 weeks after the progressed EDSS assessment time point (Figure 2). Using a roving EDSS score rather than the fixed study baseline EDSS score as reference reduced the potential bias toward the selection of entirely relapse-free patients when progression events unrelated to relapse were analyzed.

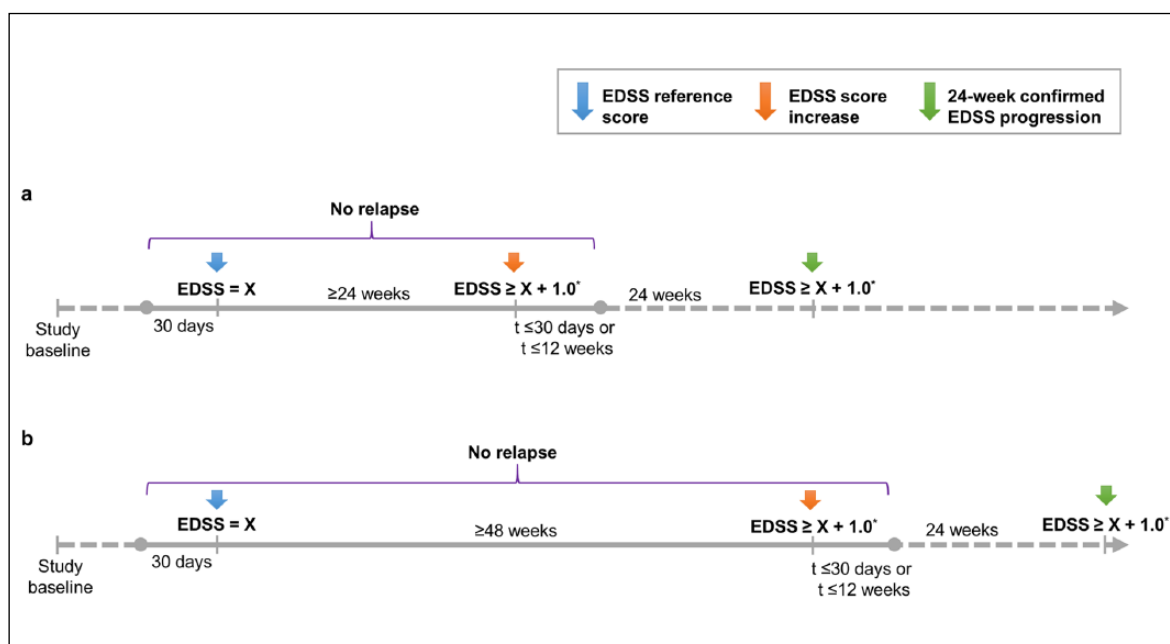
### Sensitivity analysis

A sensitivity analysis of the roving EDSS reference system was performed with progression events unrelated to relapse (increases in EDSS score of  $\geq 1.5$  points from an EDSS score of 0.0,  $\geq 1.0$  point from an EDSS score of 1.0–5.5, or  $\geq 0.5$  point from an

EDSS score  $\geq 6.0$ ) excluded if recorded in reference to an EDSS score that was both lower than study baseline EDSS score and not confirmed after  $\geq 12$  weeks. (For example, if a recorded EDSS score was lower than the score at study baseline, then, to qualify for use as a progression reference, this EDSS score must also have been confirmed by a second EDSS score at least as low as the first score  $\geq 12$  weeks later.)

### Statistical analysis

Baseline characteristics are presented using summary statistics as appropriate. Cumulative probabilities of EDSS worsening and progression unrelated to relapse were estimated using the Kaplan–Meier method. If the onset of EDSS worsening or progression unrelated to relapse occurred  $\geq 12$  weeks after the last dose of natalizumab, this event was excluded and the patient was censored 12 weeks after the last dose. However, confirmation of the worsening or



**Figure 2.** Schematic of methodological assessment of Expanded Disability Status Scale (EDSS) progression unrelated to relapses confirmed (a)  $\geq 24$  or (b)  $\geq 48$  weeks apart using a roving EDSS reference.

*t*: time following EDSS increase with no concurrent relapse.

\*Increase in EDSS score of  $\geq 1.5$  points from a score of 0.0,  $\geq 1.0$  point from a score of 1.0–5.5, or  $\geq 0.5$  points from a score  $\geq 6.0$ .

progression event could occur  $\geq 12$  weeks after the last dose of natalizumab. The analysis allowed multiple events of progression; when a patient had  $\geq 2$  confirmed progression events, the event with the earliest onset date was used. Patients who dropped out of the study before week 288 and who had not had any progression events were censored 12 weeks after the last dose or at week 288, whichever was earlier.

Cumulative probabilities of EDSS worsening and progression unrelated to relapse using a roving EDSS reference were also analyzed in subgroups based on age at baseline ( $\leq 37$  years and  $> 37$  years), baseline EDSS score ( $\leq 3.5$  and  $> 3.5$ ), the number of relapses prior to starting natalizumab ( $< 2$  and  $\geq 2$ ), and MS disease duration at baseline ( $\leq 7$  years and  $> 7$  years) in which patients were stratified according to their distribution above or below the median value at study baseline. Analyses were conducted using SAS/STAT software (version 9.3, SAS Institute, Cary, NC, USA).

## Results

As of 1 May 2014, a total of 5623 patients were enrolled in TOP. Of these patients, 5562 had baseline and follow-up EDSS scores and were included in the analysis (Table 1). At the time of data extraction, patients had been on natalizumab treatment for a

median (25th percentile, 75th percentile) of 108.3 (57.4, 176.6) weeks.

### *Disability worsening and progression unrelated to relapse using study baseline EDSS score as a fixed reference*

When the baseline EDSS score was used as a fixed reference point, the cumulative probabilities of 24-week-confirmed EDSS worsening  $\geq 24$  and  $\geq 48$  weeks apart (i.e. with  $\geq 24$  or  $\geq 48$  weeks between the reference and the EDSS score increase) were similar (20.3% and 19.5%, respectively; Figure 3(a) and (b); Table 2). Exclusion of relapse-associated events reduced the number of confirmed progression events by approximately 50%. Similar results were observed when the 30-day or 12-week relapse cutoff was used to ensure that the captured progression events did not reflect an emerging subsequent relapse.

### *Disability worsening and progression unrelated to relapse using a roving EDSS reference value*

When a roving EDSS reference value was used, the cumulative probability at 288 weeks in TOP of 24-week-confirmed EDSS worsening between EDSS assessments  $\geq 24$  weeks apart (37.1%; Figure 3(c))

**Table 1.** Baseline characteristics.

| Characteristics                                            | TOP patients ( <i>n</i> = 5562) |
|------------------------------------------------------------|---------------------------------|
| Age, years                                                 |                                 |
| Mean (SD)                                                  | 37.1 (9.73)                     |
| Median (min, max)                                          | 37.0 (12, 70)                   |
| Gender, <i>n</i> (%)                                       |                                 |
| Male                                                       | 1556 (28.0)                     |
| Female                                                     | 4006 (72.0)                     |
| MS duration, years ( <i>n</i> = 5545)                      |                                 |
| Mean (SD)                                                  | 8.61 (6.687)                    |
| Median (min, max)                                          | 7.15 (0.0, 43.9)                |
| Baseline EDSS score ( <i>n</i> = 5555)                     |                                 |
| Mean (SD)                                                  | 3.45 (1.629)                    |
| Median (range)                                             | 3.5 (0.0, 9.5)                  |
| Prior DMTs, <i>n</i> (%)                                   |                                 |
| 0                                                          | 538 (9.7)                       |
| 1                                                          | 2506 (45.1)                     |
| $\geq 2$                                                   | 2518 (45.3)                     |
| Prior relapses in last year ( <i>n</i> = 5561)             |                                 |
| Mean (SD)                                                  | 2.0 (1.01)                      |
| Median (min, max)                                          | 2 (0, 10)                       |
| Natalizumab doses received before enrollment, <i>n</i> (%) |                                 |
| 0                                                          | 2374 (42.7)                     |
| 1                                                          | 1278 (23.0)                     |
| 2                                                          | 1010 (18.2)                     |
| 3                                                          | 900 (16.2)                      |

DMT: disease-modifying therapy; EDSS: Expanded Disability Status Scale; MS: multiple sclerosis; SD: standard deviation.

was similar to that between assessments  $\geq 48$  weeks apart (34.9%; Figure 3(d); Table 2).

Events of confirmed disability progression unrelated to relapse (Figure 3; Table 2) represented 61%–66% of overall confirmed disability worsening. Furthermore, 2.4 and 2.2 times more progression events unrelated to relapse measured  $\geq 24$  weeks apart and  $\geq 48$  weeks apart, respectively, were captured using the roving EDSS value rather than the study baseline EDSS score as a reference (Table 2).

A sensitivity analysis using a roving EDSS reference value that required confirmation of the new reference if it was lower than the study baseline EDSS score reduced the overall number of identified worsening events by 17%–20% and reduced the number of progression events unrelated to relapse by 22%–27%. However, the analysis using the roving EDSS reference still detected approximately 50% more overall worsening events and approximately 70% more progression events unrelated to relapse than analyses

using the study baseline EDSS score as a fixed EDSS reference (Figure 4; Table 2).

*Confirmed EDSS worsening and progression unrelated to relapse stratified by baseline characteristics using a roving EDSS reference value*

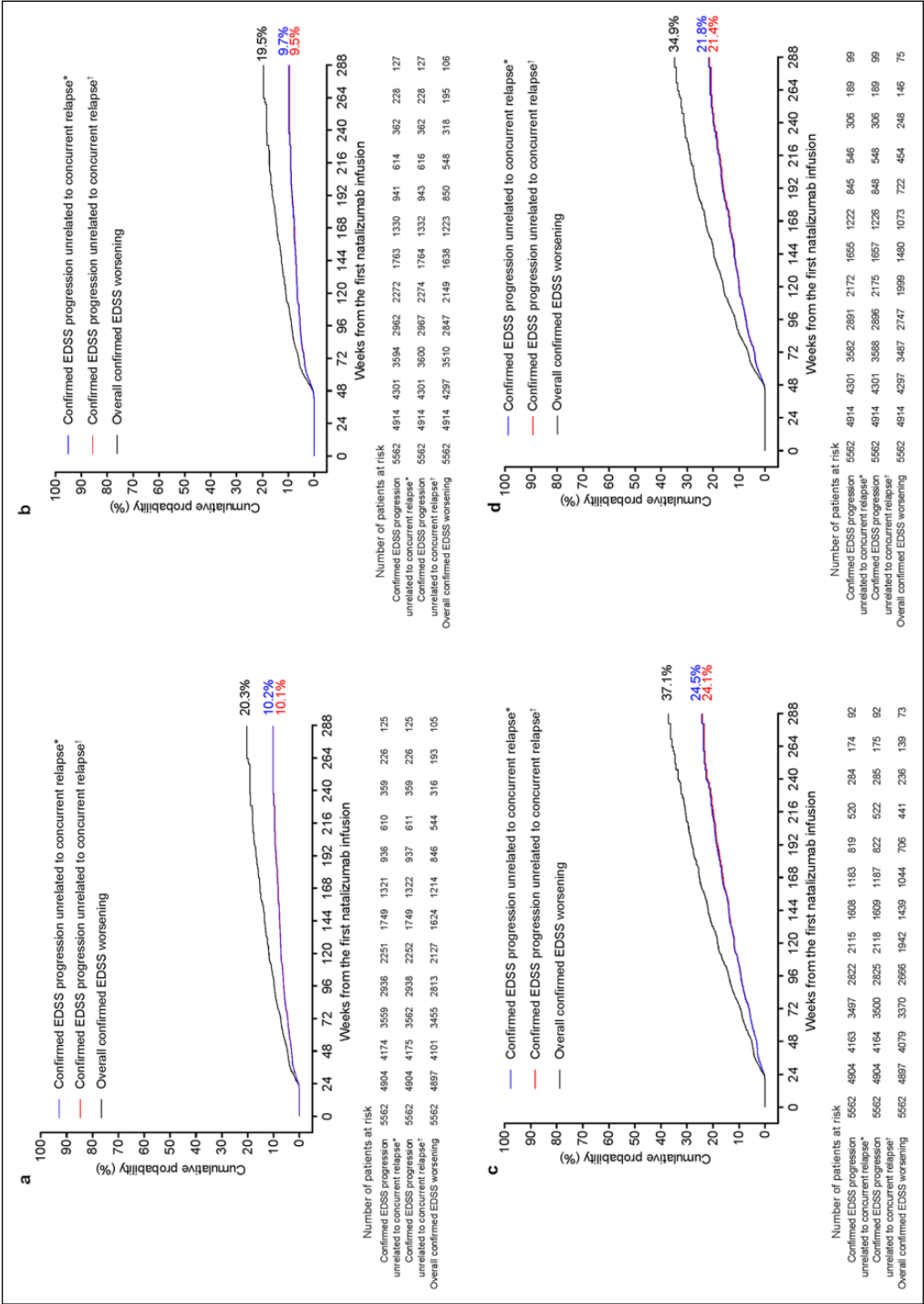
Patients in the TOP study population above the median age of 37 years at baseline had more EDSS worsening and progression events unrelated to relapse than patients  $\leq 37$  years old at baseline (Table 3); analyses of both  $\geq 24$  weeks and  $\geq 48$  weeks apart (using both the primary analysis and the more stringent sensitivity analysis) showed approximately 1.6 to 2.0 times more progression events unrelated to relapse in older than in younger patients (Table 3).

Patients who initiated natalizumab after  $< 2$  relapses in the prior year had slightly higher rates of EDSS worsening and progression unrelated to relapse than patients with  $\geq 2$  relapses (Table 3). EDSS progression events unrelated to relapse occurred at a similar rate in patients with baseline EDSS  $\leq 3.5$  (the median score) and with baseline EDSS  $> 3.5$  in both the primary and sensitivity analyses (Table 3). EDSS worsening and progression unrelated to relapse events were also similar in patients with different MS disease durations at baseline ( $\leq 7$  years or  $> 7$  years; Table 3).

## Discussion

In this study, the use of a roving EDSS reference captured more than twice as many EDSS worsening events as analyses using the study baseline EDSS score as a fixed reference. Assessment of changes in EDSS score based on a roving EDSS reference value rather than a conventional fixed study baseline EDSS reference may therefore serve as a more sensitive measure to capture events of disability progression in clinical trials and long-term observational MS studies. To address the potential variability of the new reference EDSS score used for the roving reference analysis, the more stringent sensitivity analysis criteria included only events using a roving EDSS reference that itself had to be confirmed when lower than the study baseline EDSS score. This sensitivity analysis still revealed approximately 50%–70% more worsening and progression events than analyses using the study baseline EDSS score as a fixed reference.

Analyses using the roving reference seem especially sensitive to the detection of events unrelated to relapses. According to the revised definitions of the clinical course of MS by Lublin *et al.*,<sup>6</sup> confirmed



**Figure 3.** Cumulative probabilities (Kaplan–Meier analysis) of 24-week-confirmed Expanded Disability Status Scale (EDSS) overall worsening or progression and of 24-week-confirmed EDSS progression unrelated to relapses identified using a conventional study baseline reference for events occurring (a)  $\geq 24$  weeks apart or (b)  $\geq 48$  weeks apart or using a moving reference for events occurring between two EDSS assessments (c)  $\geq 24$  weeks apart or (d)  $\geq 48$  weeks apart.  
\*Defined as a relapse that was recorded from  $\leq 30$  days prior to the reference EDSS assessment to  $\leq 30$  days post progression assessment.  
†Defined as a relapse that was recorded from  $\leq 30$  days prior to the reference EDSS assessment to  $\leq 12$  weeks post progression assessment.

**Table 2.** Cumulative probabilities (Kaplan–Meier analysis) at 288 weeks of 24-week-confirmed EDSS worsening or progression unrelated to relapse using a fixed study baseline or roving EDSS reference value ( $n=5562$ ).

| Cumulative probability, % (95% CI)                                                                                                   | Overall confirmed EDSS worsening | Confirmed EDSS progression unrelated to relapse <sup>a</sup> |                  |
|--------------------------------------------------------------------------------------------------------------------------------------|----------------------------------|--------------------------------------------------------------|------------------|
|                                                                                                                                      |                                  | ≤30 days                                                     | ≤12 weeks        |
| Fixed study baseline EDSS reference                                                                                                  |                                  |                                                              |                  |
| EDSS assessments ≥24 weeks apart                                                                                                     | 20.3 (18.0–22.5)                 | 10.2 (8.9–11.6)                                              | 10.1 (8.7–11.4)  |
| EDSS assessments ≥48 weeks apart                                                                                                     | 19.5 (17.3–21.7)                 | 9.7 (8.3–11.0)                                               | 9.5 (8.1–10.8)   |
| Roving EDSS reference value                                                                                                          |                                  |                                                              |                  |
| EDSS assessments ≥24 weeks apart                                                                                                     | 37.1 (33.5–40.5)                 | 24.5 (21.6–27.3)                                             | 24.1 (21.2–26.9) |
| EDSS assessments ≥48 weeks apart                                                                                                     | 34.9 (31.5–38.3)                 | 21.8 (19.1–24.5)                                             | 21.4 (18.7–24.2) |
| Roving EDSS reference value (sensitivity analysis)                                                                                   |                                  |                                                              |                  |
| EDSS assessments ≥24 weeks apart                                                                                                     | 29.7 (26.4–32.9)                 | 17.9 (15.4–20.3)                                             | 17.6 (15.1–20.0) |
| EDSS assessments ≥48 weeks apart                                                                                                     | 28.9 (25.7–32.2)                 | 17.1 (14.7–19.6)                                             | 16.8 (14.3–19.2) |
| EDSS: Expanded Disability Status Scale; CI: confidence interval.                                                                     |                                  |                                                              |                  |
| <sup>a</sup> With no concurrent relapse from 30 days prior to reference score until indicated time after the increase in EDSS score. |                                  |                                                              |                  |

EDSS: Expanded Disability Status Scale; CI: confidence interval.

<sup>a</sup>With no concurrent relapse from 30 days prior to reference score until indicated time after the increase in EDSS score.

disability progression unrelated to relapse (measured here using a roving window of EDSS assessment over an approximate 1-year period) may represent a reliable clinical diagnostic signature for SPMS. A defining feature of SPMS is ongoing disability progression, and our analysis assesses only the first on-study progression event. It is possible that “disability progression unrelated to relapse” as identified here was in fact due to a subclinical relapse in some patients. However, because the current definition of SPMS includes patients with or without occasional relapses,<sup>6</sup> patients with progression occurring over ≥24 weeks and even more so over ≥48 weeks would meet the current SPMS criteria. In parallel to the conventional definition of SPMS phenotype, such time-to-event analysis of confirmed disability progression unrelated to relapse could lead to a more sensitive and specific metric-based evaluation of SPMS onset and/or SPMS disease course.

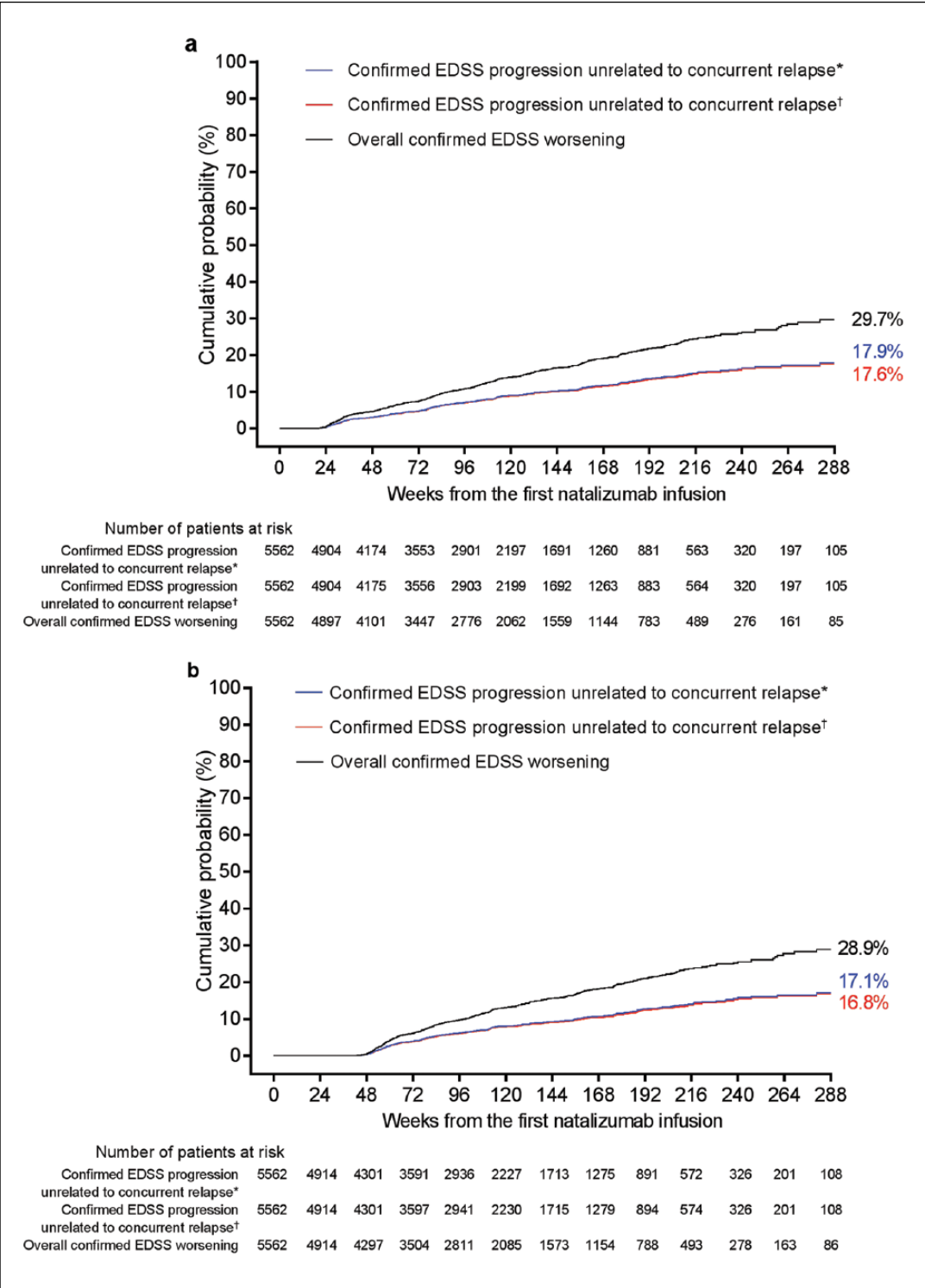
The greater sensitivity of this assessment could allow detection of treatment effects in sample sizes that are smaller than those typically needed to assess disability progression. In this work, the sensitivity to detect events of EDSS worsening and progression unrelated to relapses refers to the greater ability of a roving EDSS methodology to identify disability change. This should be distinguished from the conventional definition of “sensitivity” used in the context of diagnostic tests, which refers more specifically to the proportion of true positives identified.<sup>17</sup>

Using a roving EDSS reference system to analyze specific baseline patient characteristics suggests that being part of the older age group but not disease

duration or baseline EDSS score increased the risk of progression. This finding is consistent with the stronger correlation of age as compared to EDSS score or disease duration with risk of SPMS onset described in previous studies.<sup>12–14</sup> The similar number of events detected when applying more stringent criteria for confirmation intervals and absence of relapse should increase confidence in the specificity of this measure.

A limitation of this analysis is that, as with any assessment of disability progression, data are missing due to missing patient visits.

Caution is necessary if our proposed approach to the assessment of disability progression events unrelated to relapses was to be applied in the setting of comparative trials in patients with RRMS because of the bias related to the influence of post-randomization factors. For instance, any post-baseline change in relapse rate and/or improvement/decrease in EDSS due to differential treatment effects would impact sensitivity to detect subsequent progression events. This potential bias would need to be taken into account in the predefined statistical analysis plan and more importantly in the interpretation of the results. Implementation of adequate marginal structural models to account for post-randomization time-varying factors and interval censoring around the occurrence of relapses could help to reduce imbalance in the cumulative epoch time in which patients are at risk of disability progression events unrelated to relapse. Such prospectively implemented measures would reduce but not completely eliminate bias. This taken into account, the use of a roving reference system to



**Figure 4.** Cumulative probabilities (Kaplan–Meier analysis) of 24-week-confirmed Expanded Disability Status Scale (EDSS) overall worsening or progression and of 24-week-confirmed EDSS progression unrelated to relapses using sensitivity analysis criteria (roving baseline confirmed at 12 weeks). Kaplan–Meier plots show the cumulative probability of events occurring between two EDSS assessments: (a)  $\geq 24$  weeks apart or (b)  $\geq 48$  weeks apart.  
\*Defined as a relapse that was recorded from  $\leq 30$  days prior to the reference EDSS assessment to  $\leq 30$  days post progression assessment.  
†Defined as a relapse that was recorded from  $\leq 30$  days prior to the reference EDSS assessment to  $\leq 12$  weeks post progression assessment.

**Table 3.** Cumulative probabilities (Kaplan–Meier analysis) at 288 weeks of confirmed EDSS worsening or progression unrelated to relapse using a roving EDSS reference value and stratified by the baseline characteristics of age, EDSS score, number of relapses in the prior year, and MS disease duration.

| Cumulative probability, % (95% CI)               | Overall confirmed EDSS worsening |                      | Confirmed EDSS progression unrelated to relapse <sup>a</sup> |                      |                  |                      |
|--------------------------------------------------|----------------------------------|----------------------|--------------------------------------------------------------|----------------------|------------------|----------------------|
|                                                  |                                  |                      | ≤30 days                                                     |                      | ≤12 weeks        |                      |
|                                                  | Primary analysis                 | Sensitivity analysis | Primary analysis                                             | Sensitivity analysis | Primary analysis | Sensitivity analysis |
| Age ≤37 years ( <i>n</i> = 2932)                 |                                  |                      |                                                              |                      |                  |                      |
| EDSS assessments ≥24 weeks apart                 | 31.8 (27.4–36.1)                 | 25.5 (21.5–29.6)     | 18.9 (15.8–22.0)                                             | 12.6 (10.3–14.9)     | 18.4 (15.4–21.5) | 12.2 (10.0–14.5)     |
| EDSS assessments ≥48 weeks apart                 | 30.2 (26.0–34.5)                 | 25.0 (21.0–29.1)     | 16.6 (13.7–19.5)                                             | 11.8 (9.5–14.1)      | 16.3 (13.4–19.2) | 11.5 (9.2–13.7)      |
| Age >37 years ( <i>n</i> = 2630)                 |                                  |                      |                                                              |                      |                  |                      |
| EDSS assessments ≥24 weeks apart                 | 42.8 (37.6–48.0)                 | 34.3 (29.3–39.3)     | 30.6 (25.9–35.3)                                             | 23.8 (19.4–28.2)     | 30.3 (25.6–35.0) | 23.5 (19.2–27.9)     |
| EDSS assessments ≥48 weeks apart                 | 39.9(34.8–45.1)                  | 33.2 (28.2–38.2)     | 27.5 (23.0–32.1)                                             | 23.1 (18.7–27.5)     | 27.2 (22.6–31.7) | 22.7 (18.3–27.1)     |
| EDSS score ≤3.5 ( <i>n</i> = 3202)               |                                  |                      |                                                              |                      |                  |                      |
| EDSS assessments ≥24 weeks apart                 | 35.3 (30.7–40.0)                 | 28.9 (24.4–33.3)     | 24.0 (20.1–27.9)                                             | 17.8 (14.3–21.2)     | 23.6 (19.7–27.4) | 17.5 (14.1–20.9)     |
| EDSS assessments ≥48 weeks apart                 | 33.4 (28.8–38.0)                 | 28.1 (23.7–32.5)     | 21.6 (17.8–25.4)                                             | 17.2 (13.8–20.7)     | 21.4 (17.6–25.2) | 17.0 (13.6–20.4)     |
| EDSS score >3.5 ( <i>n</i> = 2353)               |                                  |                      |                                                              |                      |                  |                      |
| EDSS assessments ≥24 weeks apart                 | 39.8 (34.4–45.2)                 | 31.1 (26.2–35.9)     | 25.3 (21.2–29.4)                                             | 18.2 (14.7–21.6)     | 25.0 (20.8–29.1) | 17.8 (14.3–21.3)     |
| EDSS assessments ≥48 weeks apart                 | 37.2 (31.9–42.4)                 | 30.3 (25.4–35.2)     | 22.1 (18.4–25.9)                                             | 17.1 (13.6–20.6)     | 21.6 (17.8–25.4) | 16.6 (13.1–20.1)     |
| Relapses in the prior year <2 ( <i>n</i> = 1959) |                                  |                      |                                                              |                      |                  |                      |
| EDSS assessments ≥24 weeks apart                 | 40.6 (34.6–46.6)                 | 32.5 (27.1–37.9)     | 28.3 (23.4–33.1)                                             | 21.6 (17.4–25.9)     | 27.8 (23.0–32.6) | 21.5 (17.2–25.7)     |
| EDSS assessments ≥48 weeks apart                 | 37.5 (31.8–43.3)                 | 31.6 (26.2–37.0)     | 25.1 (20.7–29.6)                                             | 20.9 (16.7–25.2)     | 24.8 (20.4–29.3) | 20.7 (16.4–24.9)     |
| Relapses in the prior year ≥2 ( <i>n</i> = 3602) |                                  |                      |                                                              |                      |                  |                      |
| EDSS assessments ≥24 weeks apart                 | 35.3 (31.1–39.6)                 | 28.1 (24.2–32.1)     | 22.5 (19.0–26.0)                                             | 15.9 (12.9–18.9)     | 22.2 (18.7–25.6) | 15.6 (12.6–18.5)     |
| EDSS assessments ≥48 weeks apart                 | 33.5 (29.3–37.7)                 | 27.5 (23.5–31.5)     | 20.1 (16.7–23.5)                                             | 15.2 (12.2–18.2)     | 19.7 (16.3–23.1) | 14.8 (11.8–17.8)     |
| MS disease duration ≤7 years ( <i>n</i> = 2732)  |                                  |                      |                                                              |                      |                  |                      |
| EDSS assessments ≥24 weeks apart                 | 35.4 (30.2–40.6)                 | 27.5 (23.0–32.0)     | 23.8 (19.1–28.6)                                             | 17.0 (13.0–21.0)     | 23.6 (18.9–28.4) | 16.8 (12.8–20.9)     |
| EDSS assessments ≥48 weeks apart                 | 32.7 (27.8–37.7)                 | 27.0 (22.5–31.6)     | 21.1 (16.6–25.6)                                             | 16.4 (12.4–20.5)     | 20.8 (16.3–25.3) | 16.2 (12.2–20.2)     |
| MS disease duration >7 years ( <i>n</i> = 2813)  |                                  |                      |                                                              |                      |                  |                      |
| EDSS assessments ≥24 weeks apart                 | 38.5 (33.8–43.2)                 | 31.6 (26.9–36.2)     | 25.1 (21.8–28.5)                                             | 18.7 (15.7–21.7)     | 24.5 (21.2–27.9) | 18.3 (15.3–21.3)     |
| EDSS assessments ≥48 weeks apart                 | 36.7 (32.0–41.4)                 | 30.6 (26.0–35.2)     | 22.6 (19.2–25.9)                                             | 17.8 (14.9–20.8)     | 22.1 (18.8–25.4) | 17.4 (14.4–20.4)     |

EDSS: Expanded Disability Status Scale; MS: multiple sclerosis; CI: confidence interval.

<sup>a</sup>With no concurrent relapse from 30 days prior to reference score until indicated time after the EDSS score increase.

more accurately capture confirmed disability worsening or progression unrelated to relapse could be applied beyond the EDSS to other clinical disability outcome measures such as the Timed 25-Foot Walk, the 9-Hole Peg Test, low contrast letter acuity, cognitive function testing,<sup>18–21</sup> or multicomponent endpoints using various combinations of the above.<sup>22</sup>

### Acknowledgements

All named authors meet the International Committee of Medical Journal Editors' (ICMJE) criteria for authorship for this manuscript and take responsibility for the integrity of the work as a whole. Biogen provided funding for medical writing support in the development of this manuscript; Edwin Thrower and Alison Adams from Ashfield Healthcare Communications (Middletown, CT), based on input from authors, wrote the first draft and revised subsequent drafts of the manuscript, and Joshua Safran from Ashfield Healthcare Communications copyedited and styled the manuscript per journal requirements. Biogen reviewed and provided feedback on the manuscript to the authors. The authors had full editorial control of the manuscript, and provided their final approval of all content.

### 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: L.K.'s Institution (University Hospital Basel) received in the last 3 years and used exclusively for research support: steering committee, advisory board, and consultancy fees from Actelion, Addex, Bayer HealthCare, Biogen, Biotica, Genzyme, Lilly, Merck, Mitsubishi, Novartis, Ono Pharma, Pfizer, Receptos, Sanofi, Santhera, Siemens, Teva, UCB, and Xenoport; speaker fees from Bayer HealthCare, Biogen, Merck, Novartis, Sanofi, and Teva; support of educational activities from Bayer HealthCare, Biogen, CSL Behring, Genzyme, Merck, Novartis, Sanofi, and Teva; license fees for Neurostatus products; and grants from Bayer HealthCare, Biogen, F. Hoffmann-La Roche Ltd, Merck, Novartis, the Swiss Multiple Sclerosis Society, the Swiss National Research Foundation, the European Union, and the Roche Research Foundation. M.T. received compensation for consulting from Biogen, Merck Serono, and Novartis; speaker honoraria from Biogen, Merck Serono, Novartis, Sanofi, and Teva; and research grants from Biogen, Merck Serono, and Novartis. H.B. received compensation for consulting from Biogen, Merck Serono, and Novartis and research support from Biogen and Merck Serono. H.W. received honoraria from Bayer, Biogen, Medac, Merck Serono, Novo Nordisk, Sanofi, Schering, and Teva; was a consultant for Bayer Vital/Schering, Biogen, Medac, Merck Serono,

Novartis, Novo Nordisk, Sanofi, and Teva; and received research support from Bayer, Biogen, Medac, Merck Serono, Novo Nordisk, Sanofi, Schering, and Teva. T.S. received honoraria for consultancy and funding for travel from Biogen and Novartis. F.P., and Q.D. are employees of and own stock and/or stock options in Biogen. S.B. and H.K. were employees of Biogen at the time this analysis was performed and are now employees of F. Hoffmann-La Roche Ltd. Y.C. was an employee of Biogen at the time this analysis was performed and is now an employee of Shire. (F. Hoffmann-La Roche Ltd and Shire were not in any way associated with this study).

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study and article were funded by Biogen.

### References

1. Meyer-Moock S, Feng YS, Maeurer M, et al. Systematic literature review and validity evaluation of the Expanded Disability Status Scale (EDSS) and the Multiple Sclerosis Functional Composite (MSFC) in patients with multiple sclerosis. *BMC Neurol* 2014; 14: 58.
2. Rudick RA, Lee JC, Cutter GR, et al. Disability progression in a clinical trial of relapsing-remitting multiple sclerosis: Eight-year follow-up. *Arch Neurol* 2010; 67: 1329–1335.
3. Kalincik T, Cutter GR, Spelman T, et al. Defining reliable disability outcomes in multiple sclerosis. *Brain* 2015; 138: 3287–3298.
4. European Medicines Agency. Guideline on clinical investigation of medicinal products for the treatment of multiple sclerosis, 2015, [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2015/03/WC500185161.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/03/WC500185161.pdf) (accessed 11 June 2015).
5. Phillips JT, Giovannoni G, Lublin FD, et al. Sustained improvement in Expanded Disability Status Scale as a new efficacy measure of neurological change in multiple sclerosis: Treatment effects with natalizumab in patients with relapsing multiple sclerosis. *Mult Scler* 2011; 17: 970–979.
6. Lublin FD, Reingold SC, Cohen JA, et al. Defining the clinical course of multiple sclerosis: The 2013 revisions. *Neurology* 2014; 83: 278–286.
7. Tremlett H, Yinshan Z and Devonshire V. Natural history of secondary-progressive multiple sclerosis. *Mult Scler* 2008; 14: 314–324.

8. Lublin FD and Reingold SC. Defining the clinical course of multiple sclerosis: Results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. *Neurology* 1996; 46: 907–911.
9. Scafari A, Neuhaus A, Daumer M, et al. Onset of secondary progressive phase and long-term evolution of multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2014; 85: 67–75.
10. Confavreux C and Vukusic S. Age at disability milestones in multiple sclerosis. *Brain* 2006; 129: 595–605.
11. Confavreux C, Vukusic S and Adeleine P. Early clinical predictors and progression of irreversible disability in multiple sclerosis: An amnesic process. *Brain* 2003; 126: 770–782.
12. Koch M, Mostert J, Heersema D, et al. Progression in multiple sclerosis: Further evidence of an age dependent process. *J Neurol Sci* 2007; 255: 35–41.
13. Scafari A, Neuhaus A, Daumer M, et al. Age and disability accumulation in multiple sclerosis. *Neurology* 2011; 77: 1246–1252.
14. Tutuncu M, Tang J, Zeid NA, et al. Onset of progressive phase is an age-dependent clinical milestone in multiple sclerosis. *Mult Scler* 2013; 19: 188–198.
15. Lorscheider J, Buzzard K, Jokubaitis V, et al. Defining secondary progressive multiple sclerosis. *Brain* 2016; 139: 2395–2405.
16. Butzkueven H, Kappos L, Pellegrini F, et al.; TYSABRI Observational Program Investigators. Efficacy and safety of natalizumab in multiple sclerosis: Interim observational programme results. *J Neurol Neurosurg Psychiatry* 2014; 85: 1190–1197.
17. Altman DG and Bland JM. Diagnostic tests. 1: Sensitivity and specificity. *BMJ* 1994; 308: 1552.
18. Bosma L, Kragt JJ, Polman CH, et al. Walking speed, rather than Expanded Disability Status Scale, relates to long-term patient-reported impact in progressive MS. *Mult Scler* 2013; 19: 326–333.
19. Bosma LV, Sonder JM, Kragt JJ, et al. Detecting clinically-relevant changes in progressive multiple sclerosis. *Mult Scler* 2015; 21: 171–179.
20. Cadavid D, Jurgensen S and Lee S. Impact of natalizumab on ambulatory improvement in secondary progressive and disabled relapsing-remitting multiple sclerosis. *PLoS ONE* 2013; 8: e53297.
21. Galetta KM and Balcer LJ. Measures of visual pathway structure and function in MS: Clinical usefulness and role for MS trials. *Mult Scler Relat Disord* 2013; 2: 172–182.
22. Zhang J, Waubant E, Cutter GR, et al. Composite end points to assess delay of disability progression by MS treatments. *Mult Scler* 2014; 20: 1494–1501.

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