2015-05

Identifying a Minimally Important Change Threshold for The Multiple Sclerosis Walking Scale-12 (MSWS-12)

Mehta, L

MULTIPLE SCLEROSIS JOURNAL

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

Identifying an Important Change Estimate for the Multiple Sclerosis Walking Scale-12 (MSWS-12v1) For Interpreting Clinical Trial Results

Lahar Mehta, MD[1]; Manjit McNeill, MSc[2]; Jeremy Hobart, FRCP, PhD[3]; Kathleen Wyrwich, PhD[4]; Jiat-Ling Poon, PhD[4]; Priscilla Auguste, MHS[4]; John Zhong, PhD[1]; Jacob Elkins, MD[1]

[1]Biogen, Cambridge, MA, USA; [2]Biogen, Maidenhead, United Kingdom;

[3]Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, UK;

[4]Evidera, Bethesda, MD, USA

Running Title: Important Change Estimate for MSWS-12v1

Keywords: important change, meaningful improvement, responder definition, MSWS-12, patient-reported outcomes, multiple sclerosis, walking ability, anchor-based, distribution-based

Corresponding Author:

Kathleen W. Wyrwich, PhD

7101 Wisconsin Avenue, Suite 1400

Bethesda, MD 20814

USA

kathy.wyrwich@evidera.com

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

+1-314-773-8000 (phone)

+1-314-773-8004 (fax)

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

ABSTRACT

Background: The 12-question Multiple Sclerosis (MS) Walking Scale (MSWS-12v1) is a widely-used patient-reported outcome (PRO) measure of walking ability in MS.

Objective: To estimate the magnitude of an important change in MSWS-12v1 scores for the interpretation of meaningful subject-level improvements across a 6-month trial of MS patients with walking disability.

Methods: MOBILE was a 6-month exploratory study assessing fampridine's effect on walking ability in 132 people with MS. Three PRO measures assessed walking ability: MSWS-12v1, EuroQol 5-Dimension-5 Level (EQ-5D-5L) mobility question, and a patient global impression of change (PGIC) in overall walking ability. Pre-specified anchor- and distribution-based analyses estimated the MSWS-12v1 change scores representing an important change for participants. Results were triangulated to propose a single best value indicating meaningful improvement.

Results: Using Baseline to Week 2 through Week 24 change scores, anchor-based analyses demonstrated mean and median improvements of 5.2-6.6 (PGIC) and 9.7-13.4 (EQ-5D-5L mobility) points on the MSWS-12v1 indicated meaningful improvements. The distribution-based estimate was 6.8 points. Triangulation across the results suggested an 8-point reduction in MSWS-12v1 score represents an important subject-level change in these participants.

Conclusion: In similar MS clinical trials, an 8-point improvement on the MSWS-12v1is a reasonable estimate of meaningful improvement in walking ability.

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

INTRODUCTION

Patient-reported outcomes (PROs) are increasingly used as efficacy endpoints in clinical trials to measure symptoms and disease impact. It is therefore useful to understand the clinical importance or meaningfulness of changes in PRO scores to the individuals receiving treatments. The concept of a *minimal clinically important difference (MCID)* was suggested to address the need for meaningful interpretation of PRO scores. Initially, the MCID was defined as the ''smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient's management'' [1]. More recently, a similar concept has been referred to as the *responder definition (RD)*, or "the individual patient PRO score change over a predetermined time period that should be interpreted as a treatment benefit" [2]. Both of these terms seek to understand the magnitude of a PRO measure's change over time that demonstrates a meaningful improvement to an individual patient.

The widely-used 12-question Multiple Sclerosis (MS) Walking Scale (MSWS-12v1) is a patient-reported measure of walking ability in MS [3]. Evidence from multiple studies supports its robust measurement performance [3-8]. With approximately 75% of people with MS identifying gait, mobility, and balance as key physical problems, the MSWS-12v1 has been used frequently in contemporary MS clinical trials to assess these important issues related to walking ability [9-12]. The objective of this current analysis was to estimate the MSWS-12v1 change score representing a meaningful improvement in walking ability, using standard methods, in data from walking-disabled people with MS participating in a 6-month, randomized treatment trial.

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

METHODS

Participants

Data from MOBILE (NCT01597297; EudraCT 2012-000368-90) are reported elsewhere [13]. Briefly, it was a multicenter, randomized, double-blind, placebo-controlled trial to assess the longer-term effects of fampridine-PR 10 mg BD tablets on self-assessed walking disability, and the impact of treatment on overall walking ability in MS over 24 weeks. A total of 132 walking-disabled people with MS, aged 18 to 70, were enrolled at 24 sites in Belgium, Canada, Italy, the Netherlands, Sweden, and the United Kingdom. Eligibility criteria were: any MS subtype, screening visit Expanded Disability Status Scale (EDSS) [14] score of 4 to 7 and clinically stable (no MS exacerbation within 60 days of screening). Participants were randomized in a 1:1 ratio on Day 1 to receive either fampridine-PR 10 mg BD or matching placebo. As shown in Figure 1, scheduled study visits were at Day 1, Weeks 2, 4, 8, 12, 16, 20, 24, and 26 (two weeks post treatment discontinuation). Unscheduled visits (if required) occurred within five days of the onset of symptoms indicating possible relapse, suspected seizure, or possible deterioration of renal function.

Measures

Multiple Sclerosis Walking Scale-12 version 1 (MSWS-12v1)

The MSWS-12v1 is a 12-item questionnaire that asks subjects to rate, on a 5-point scale (from 1=not at all to 5=extremely), their MS-related mobility limitations during the preceding 2 weeks. The MSWS-12v1 was completed at Screening and all scheduled and unscheduled study visits. To calculate the MSWS-12v1 score, the sum of the 12 questions (sum range: 12-60) is

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

transformed to have a range of 0 to 100, where higher scores indicate greater walking limitations. For individuals completing all 12 items at each study visit, the possible score changes over time increase or decrease in increments of ~2 points (also known as a PRO *state change*) [15] on the 0-100 point scale. For visits where responses to ≤6 of 12 MSWS-12v1 questions were missing, the person-specific mean score from the answered questions was imputed as the score for each missing question. For visits where ≥7 of the 12 component questions were not answered, the MSWS-12v1 score was considered missing [3].

The MSWS-12v1 includes an initial screening question asking subjects if they are unable to walk. People responding "Yes" were instructed not to complete the questionnaire. Their MSWS-12v1 score was set to 100 if none of the MSWS-12v1 questions were completed; however, few subjects (n =3) replied in this manner during the study. Irrespective of the response to this initial question, if subjects responded to the 12 questions, these were used in deriving the MSWS-12v1 score.

Patient Global Impression of Change (PGIC) in Walking

A Patient Global Impression of Change (PGIC) assessment of walking ability is a single question with multiple response options. The PGIC used in the MOBILE trial asked: "In the past 7 days, how much has the study drug affected your overall walking?" Response options were: 1=very much worse, 2=much worse, 3=slightly worse, 4=unchanged, 5=slightly improved, 6=much improved, 7=very much improved. Participants completed the PGIC at Weeks 2, 4, 8, 12, 16, 20, and 24.

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

EuroQoL 5-Dimension 5-Level (EQ-5D-5L)

The EQ-5D-5L is a generic self-reported measure of health status [16]. It includes five questions concerning mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. A summary utility index value can be computed from subject's response to these five questions [17]; values range from -0.594 (worst possible health score) to 0 (death) to 1.000 (best health). Due to the direct relevance of the response options to walking ability, only the single EQ-5D-5L domain question measuring mobility (five-level response options: 1=no problems in walking about, 2=slight problems in walking about, 3=moderate problems in walking about, 4=severe problems in walking about, or 5=unable to walk about) was used as an anchor to estimate an important change in MSWS-12v1 score. The EQ-5D-5L was administered on Day 1, and at Weeks 4, 8, 12, 16, 20, and 24.

Multiple Sclerosis Impact Scale version 1 (MSIS-29v1)

The MSIS-29v1 is a 29-item questionnaire measuring the physical (20 items) and psychological (9 items) impact of MS. Each item has 5 response options scored from 1=not at all limited/bothered to 5=extremely limited/bothered. MSIS-29v1 subscale scores (physical or psychological) are calculated by transforming the sum of the subscale items to a score ranging from 0 (no impact of MS) to 100 (extreme impact of MS). The MSIS-29v1 was completed at Screening, and at each scheduled and unscheduled visit.

Statistical Analyses

All analyses were pre-specified. Both anchor- and distribution-based methods were used to inform the important change estimate for the MSWS-12v1 score associated with a meaningful

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

improvement in walking ability during the MOBILE study. Negative MSWS-12v1 change scores indicate improvements in walking.

Anchor-based Analyses

MSWS-12v1 change scores were correlated with the proposed anchors that directly assessed walking ability (PGIC scores and EQ-5D-5L mobility item change scores) to determine if they met the suggestion that associations should exceed 0.30-0.35 [18].

PGIC

The PGIC was used as one anchor to estimate the MSWS-12v1 score change associated with a meaningful patient-reported change in walking. The retrospective PGIC responses of '5=slightly improved' were selected to represent time points where participants registered meaningful changes in walking ability. Each time a participant gave a PGIC score of '5=slightly improved' we computed the MSWS-12v1 change score from the preceding visit. As the PGIC was collected at seven visits, each participant could contribute between zero and seven MSWS-12v1 change scores. To ensure that each participant contributed only one MSWS-12v1 change score to the group-level analyses we generated a single value for participants who gave at least one PGIC responses of '5=slightly improved'. Specifically, we computed the mean (Method 1) and median (Method 2) of the individual's MSWS-12v1 change scores. Group level MSWS-12v1 change scores were then computed (Method 1 = mean [SD], range; Method 2 = median [IQR], range).

EQ-5D-5L Mobility Question

As a second anchor-based approach, the MSWS-12v2 change scores associated with meaningful changes in cross-sectional EQ-5D-5L mobility question were examined. The EQ-5D-5L was

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

collected at seven time points, giving each participant up to six EQ-5D-5L change scores, computed from baseline (e.g., EQ-5D-5L score at Week 4 minus EQ-5D-5L score at Day 1 yielded one EQ-5D-5L change score). For each participant, the median EQ-5D-5L change score across visits was computed. Participants with a median EQ-5D-5L change score = 1 were considered to have reported an important change. For these participants, the mean MSWS-12v1 change scores across visits were determined. Finally, across participants with median EQ-5D-5L change scores = 1 (one level of improvement on the mobility scale), mean and median of the mean MSWS-12v1 change scores were computed.

Distribution-based Approach

The standard error of measurement (SEM) was calculated as the distribution-based estimate of meaningful change for the MSWS-12v1 [2]. Using data from Screening and Day 1, the SEM was calculated as SEM= $\sigma_x\sqrt{1-r_{xx}}$ [19], where σ_x is the MSWS-12v1 baseline score standard deviation (Baseline = mean of Screening and Day 1 scores), and r_{xx} is the MSWS-12v1 reliability. Test-retest reliability was estimated via the intra-class correlation coefficient [ICC; method (2,1)] [20] of the MSWS-12v1 between Screening and Day 1 for subjects who remained stable between these two time points [21]. Stable subjects were identified using changes on the MSIS-29v1 physical subscale (-7.5 < MSIS-29v1 physical subscale change < 7.5) [22, 23]. In the context of MS and walking ability, the Screening to Day 1 test-retest time period of ≤14 days is reasonable for assessing test-retest reliability in walking ability in MS [24].

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

Triangulation

A single overall estimate of a meaningful change in MSWS-12v1 score was achieved by triangulating findings of the anchor- and distribution-based analyses [18, 25]. Triangulation had three stages. First, we examined the range of anchor- and distribution-based estimates. Second, we examined all MSWS-12v1 change scores for all individual participants at all visits from Week 2 to Week 24. Third, we looked for gaps (spaces) in the observed change scores in the range identified in stage one.

This process helps to identify potential change thresholds. For example, there could be naturally occurring change thresholds in the data, such as areas in the change score range where there are no change scores. There is also the possibility of naturally-occurring splits that create candidates for cut-point.

All analyses were pre-specified and performed using SAS® v9.3.

RESULTS

Table 1 shows the MOBILE sample baseline characteristics. There were 132 participants randomized to fampridine (n=68, 52%) or placebo (n=64; 48%). The mean age was 49.8 years (SD=9.0), with 54% female. At Baseline, the mean MSWS-12v1 score across all patients was 73.7, with a slightly lower mean score in the treatment group vs. placebo (71.7 vs. 75.9), with similar standard deviations. The baseline EQ-5D-5L scores demonstrated that both the baseline mean and median perceived health status were very near the midpoint between 1.00 (full health)

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

and 0 (death). Only 8 subjects (6%) had any missing data on the MSWS-12v1, each with only one missing item, and hence, limiting the impact of data imputations.

Anchor-based Approach

Correlations between MSWS-12v1 change scores and proposed anchors (PGIC, r=-0.45; EQ-5D-5L mobility question, r=-0.42) exceeded the suggested level of 0.30-0.35 [18]. This provided support for using these anchors to estimate meaningful MSWS-12v1 change scores.

Table 2 displays the MSWS-12v1 change scores (mean, SD, median, and inter-quartile ranges [IQR]) associated with each PGIC category. The number of subjects with at least one response in their Week 2 through Week 24 data in each of these PGIC response categories is provided in the second column of the table. Subjects with more than one response in any PGIC category had a single MSWS-12v1 change score data point computed for that category: either the mean (Method 1) or the median (Method 2) of their MSWS-12v1 change scores from the prior visit for the PGIC category (Table 2).

Eighty-one subjects reported, at least once, that they had 'slightly improved' since the previous study visit (51 subjects had two or more 'slightly improved' ratings). Using Method 1, the associated mean and median change (reduction) on the MSWS-12v1 for these 'slightly improved' subjects was 6.6 and 6.3 points. Under Method 2, the mean and median MSWS-12v1 change scores were 6.4 and 5.2 point reductions, respectively.

The second anchor was change on the EQ-5D-5L mobility question. Twenty-one subjects had a median 1-level improvement on the EQ-5D-5L mobility question over all study visits where both

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

EQ-5D-5L and MSWS-12v1 data were collected. Mean and median change scores for this improved group were -13.4 and -9.7, respectively. Due to the small number of individuals with this level of improvement, the median MSWS-12v1 change of -9.7 points was deemed the most appropriate meaningful change estimate using this anchor.

Distribution-based Approach

The standard deviation for the MSWS-12v1 at Baseline (19.6) and reliability calculated using MSWS-12v1 test-retest data from stable subjects (n=62; ICC=0.88) were used to calculate the SEM of the MSWS-12v1. The SEM estimate for the MSWS-12v1 was 6.8 points.

Triangulation

In order to propose a single estimate to represent a clinically meaningful improvement in MSWS-12v1 score, the values obtained from both anchor-based and distribution-based approaches were compared (Table 3). In addition, individual change scores in the MOBILE study across Week 2 through Week 24 were examined for notable gaps in this relevant range (-5.2 to -9.7); the largest of the observed gaps was between the change score values of -7.9 and -8.8 points. Based on this triangulation process [18, 25], an 8-point reduction, expressed as an integer, was selected as a reasonable estimate of a meaningful subject-level improvement on the MSWS-12v1 for this sample.

DISCUSSION

The goal of this study was to identify a meaningful change estimate for the MSWS-12v1 to guide the interpretation of subject-level changes in MS patients with walking disability. Based on these

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

results, that incorporated the complementary use of both anchor- and distribution-based methods of estimation, 8 points was estimated to be a meaningful individual improvement for the MSWS-12v1 over a 24 week period. This study was unique in that important change from Baseline was investigated using PRO anchors and associated MSWS-12v1 change scores across all visits (Week 2 through Week 24) where a relevant change could be demonstrated.

It is important to note that this estimate of 8-points change can be used for determining whether individuals with MS might be considered as having experienced a meaningful change in the same context of use (target population, clinical trial design, etc.). Applying this estimate then allows a comparison of the proportions of people achieving this threshold in different study arms. It is just as important to note that the 8-point change estimate should not be interpreted as the criterion for meaningful treatment difference for group mean change comparisons. Indeed, meaningful group mean change differences are often smaller than the meaningful individual person differences [26]. Finally, this 8-point estimate of meaningful individual response differs from the definition of responders used to identify timed walk responders in the fampridine studies [9].

Previous studies have examined important changes thresholds for the MSWS-12v1 [27-29]. One report investigated, *post hoc*, whether the difference of 6.9 MSWS-12v1 points, between timed-walk responders and non-responders observed in the pivotal fampridine trials, satisfied criteria as clinically significant [27]. Data from the two phase three trials, and multiple other studies, were examined using multiple anchor and distribution-based methods. Estimates of meaningful

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

changes varied from as little as 4 points (using subject-reported change anchors) to as much as 21 points (comparing MSWS-12v1 mean score differences of subjects who reported that they walked unaided to subject who reported that they walked with aids). Results were triangulated and suggested that individual-person change scores of <4 points were not significant, > 6 points met criteria as significant, and 4-6 points might be considered as borderline. These individual-person level estimates were then used to assist the interpretation that the 6.9 point mean group change was meaningful [27]. Others have inferred that this report ascertained MCID values of 4-6 points for MSWS-12v1 [28].

In a recent study, Baert et al. [29] investigated clinically important improvements on five walking measures, including the MSWS-12v1, in 290 people with MS from 17 European rehabilitation centers [29]. The estimation methods included: anchor-based approaches incorporating a PGIC and a therapist's global rating of change scale (GRS); and distribution-based analyses of two responsiveness indices, the smallest real change (SRC) parameter calculated at the individual (SRC_{ind}) and group (SRC_{group}) level. The MSWS-12v1 clinically important improvement estimates ranged from 10.4 to 14.1 points, and varied across subgroups.

The 8-point result in the MOBILE study is focused on a specific clinical trial population, and multiple study visits where a fast-acting medication may demonstrate a notable treatment effect at the subject level in walking ability. The contrast between this estimate of meaningful change and the Baert et al. [29] results may be due to several factors, including: 1) differing clinical enrollment characteristics (e.g., EDSS $\leq$ 6.5 vs. EDSS between 4 and 7), study design, and

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

setting (rehab interventional study vs. Phase 2 clinical trial); and 2) key differences in methodological approaches, including: the problematic use of regression analyses [30] to examine MSWS-12v1 change at specific PGIC and GRS levels; and the use of responsiveness indices [31] that incorporated a statistically higher magnitude of change than the SEM [21].

This investigation of the MSWS-12v1 meaningful change scores using the MOBILE data has several valuable features. Both anchors used in this study were PROs assessing walking ability; incorporating other patient-reported, clinical or performance-based outcomes as anchors may provide different results. From a clinical relevance perspective, this study's specific PRO anchors (PGIC and EQ-5D-5L mobility question) are directly interpretable for understanding MSWS-12v1 changes as they directly assess the impact of walking disability from the patient's perspective.

This study has a number of limitations. The sample from which the EQ-5D-based meaningful change estimates were derived was small. The two walking ability anchors generated different estimates for meaningful change, implying that the choice of anchor influences the meaningful change estimate. Both of these findings are to be expected and served to highlight complexities associated with estimating meaningful changes.

In conclusion, this study used established anchor- and distribution-based methods to derive an estimate of meaningful change for the MSWS-12v1 using data obtained from a MS clinical trial. Results suggest that an 8-point improvement in MSWS-12v1 is a reasonable estimate of an individual-person meaningful improvement in walking disability. Further studies should

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

investigate this estimate's reproducibility and applicability to aid the interpretation of other MS clinical trials.

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

CONFLICT OF INTEREST STATEMENT

Drs. Mehta, Elkins, and Zhong, and Ms. McNeill have received personal compensation for activities with Biogen as employees, and hold stock and/or stock options in Biogen. Dr. Hobart has consulted for Biogen and received support for travel, research, and clinical service development from Biogen. Drs. Wyrwich and Poon, and Ms. Auguste have received research support from Biogen.

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

FUNDING ACKNOWLEDGEMENT

This work was supported by funding from Biogen.

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

REFERENCES

1. Jaeschke R, Singer J and Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Controlled clinical trials*. 1989; 10: 407-15.

2. Food and Drug Administration. Guidance for Industry on Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. *Federal Register*. 2009; 74: 65132-3.

3. Hobart JC, Riazi A, Lamping DL, Fitzpatrick R and Thompson AJ. Measuring the impact of MS on walking ability: the 12-Item MS Walking Scale (MSWS-12). *Neurology*. 2003; 60: 31-6.

4. McGuigan C and Hutchinson M. Confirming the validity and responsiveness of the Multiple Sclerosis Walking Scale-12 (MSWS-12). *Neurology*. 2004; 62: 2103-5.

5. Motl RW and Snook EM. Confirmation and extension of the validity of the Multiple Sclerosis Walking Scale-12 (MSWS-12). *Journal of the neurological sciences*. 2008; 268: 69-73.

6. Motl RW, McAuley E and Mullen S. Longitudinal measurement invariance of the Multiple Sclerosis Walking Scale-12. *Journal of the neurological sciences*. 2011; 305: 75-9.

7. Kieseier BC and Pozzilli C. Assessing walking disability in multiple sclerosis. *Multiple sclerosis*. 2012; 18: 914-24.

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

8. Pilutti LA, Dlugonski D, Sandroff BM, et al. Further validation of multiple sclerosis walking scale-12 scores based on spatiotemporal gait parameters. *Archives of physical medicine and rehabilitation*. 2013; 94: 575-8.

9. Goodman AD, Brown TR, Krupp LB, et al. Sustained-release oral fampridine in multiple sclerosis: a randomised, double-blind, controlled trial. *Lancet*. 2009; 373: 732-8.

10. Goodman AD, Brown TR, Edwards KR, et al. A phase 3 trial of extended release oral dalfampridine in multiple sclerosis. *Annals of neurology*. 2010; 68: 494-502.

11. Zajicek JP, Hobart JC, Slade A, Barnes D, Mattison PG and Group MR. Multiple sclerosis and extract of cannabis: results of the MUSEC trial. *Journal of neurology, neurosurgery, and psychiatry*. 2012; 83: 1125-32.

12. Zajicek J, Ball S, Wright D, et al. Effect of dronabinol on progression in progressive multiple sclerosis (CUPID): a randomised, placebo-controlled trial. *The Lancet Neurology*. 2013; 12: 857-65.

13. Hupperts R, Lycke J, Short C, Gasperini C, McNeill M, Medori R, Tofil-Kaluza A, Hovenden M, Mehta L, Elkins J. Prolonged-release fampridine and walking and balance in MS: randomized controlled MOBILE trial. Mult Scler. 2015 Apr 28. pii: 1352458515581436. .

14. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*. 1983; 33: 1444-52.

15. Wyrwich KW, Spertus JA, Kroenke K, et al. Clinically important differences in health status for patients with heart disease: an expert consensus panel report. *American heart journal*. 2004; 147: 615-22.

16. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health policy*. 1990; 16: 199-208.

17. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011; 20: 1727-36.

18. Leidy NK and Wyrwich KW. Bridging the gap: using triangulation methodology to estimate minimal clinically important differences (MCIDs). *Copd*. 2005; 2: 157-65.

19. Nunnally JC and Bernstein IH. *Psychometric Theory*. 3rd ed. New York: McGraw-Hill, 1994.

20. Shrout PE and Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin*. 1979; 86: 420-8.

21. Wyrwich KW. Minimal important difference thresholds and the standard error of measurement: is there a connection? *Journal of biopharmaceutical statistics*. 2004; 14: 97-110.

22. Costelloe L, O'Rourke K, Kearney H, et al. The patient knows best: significant change in the physical component of the Multiple Sclerosis Impact Scale (MSIS-29 physical). *Journal of neurology, neurosurgery, and psychiatry*. 2007; 78: 841-4.

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

23. Phillips GA, Wyrwich KW, Guo S, et al. Responder definition of the Multiple Sclerosis Impact Scale physical impact subscale for patients with physical worsening. *Multiple sclerosis*. 2014.

24. Busse ME, Pearson OR, Van Deursen R and Wiles CM. Quantified measurement of activity provides insight into motor function and recovery in neurological disease. *Journal of neurology, neurosurgery, and psychiatry*. 2004; 75: 884-8.

25. Revicki D, Hays RD, Cella D and Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of clinical epidemiology*. 2008; 61: 102-9.

26. Wyrwich KW, Norquist JM, Lenderking WR, Acaster S and Industry Advisory Committee of International Society for Quality of Life Research. Methods for interpreting change over time in patient-reported outcome measures. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*. 2013; 22: 475-83.

27. Hobart J. Prolonged-release fampridine for multiple sclerosis: was the effect on walking ability clinically significant? *26th Congress of the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS) & 15th Annual Conference of Rehabilitation in MS (RIMS)*. Gothenburg, Sweden2010.

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

28. Motl RW, Learmonth YC, Pilutti LA, Dlugonski D and Klaren R. Validity of Minimal Clinically Important Difference Values for the Multiple Sclerosis Walking Scale-12? *European neurology*. 2014; 71: 196-202.

29. Baert I, Freeman J, Smedal T, et al. Responsiveness and Clinically Meaningful Improvement, According to Disability Level, of Five Walking Measures After Rehabilitation in Multiple Sclerosis: A European Multicenter Study. *Neurorehabilitation and neural repair*. 2014.

30. Fayers PM and Hays RD. Don't middle your MIDs: regression to the mean shrinks estimates of minimally important differences. *Qual Life Res*. 2014; 23: 1-4.

31. Pfennings LE, van der Ploeg HM, Cohen L and Polman CH. A comparison of responsiveness indices in multiple sclerosis patients. *Qual Life Res*. 1999; 8: 481-9.

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

Table 1. Baseline Characteristics of Study Population

	Placebo (N=64)	Fampridine-PR 10 mg BID (N=68)	Total (N=132)
Mean age, years (SD)	49.8 (9.3)	49.8 (8.7)	49.8 (9.0)
Male, N (%)	31 (48)	30 (44)	61 (46)
Race, N (%)			
White	63 (98)	66 (97)	129 (98)
Asian	0	2 (3)	2 (2)
Other	1 (2)	0	1 (<1)
BMI, kg/m^2: Mean (SD)	26.5 (6.2)	26.8 (4.9)	26.6 (5.6)
Expanded Disability Status Scale (EDSS): Mean (SD)	5.8 (0.9)	5.6 (0.9)	5.7 (0.9)
MSWS-12v1 Baseline score*			
Mean (SD)	75.9 (19.8)	71.7 (19.3)	73.7 (19.6)
Median (Q1, Q3)	81.3 (65.6, 90.6)	75.0 (64.1, 84.9)	78.6 (64.6, 88.0)
Min, Max	(8.3, 100)	(25.0, 100)	(8.3, 100)
MSIS-29v1 Physical Subscale Baseline score*			
Mean (SD)	53.0 (19.1)	50.9 (19.4)	51.9 (19.2)
Median (Q1, Q3)	57.5 (40.3, 65.6)	50.0 (38.1, 67.2)	52.5 (38.8, 66.3)
Min, Max	(13.1, 91.9)	(8.1, 100)	(8.1, 100)
MSIS-29v1 Psychological Subscale Baseline score*			
Mean (SD)	36.3 (20.0)	36.0 (22.2)	36.2 (21.1)
Median (Q1, Q3)	34.0 (22.9, 47.2)	32.6 (18.1, 50.7)	33.3 (20.8, 49.3)
Min, Max	(0.0, 93.1)	(1.4, 90.3)	(0.0, 93.1)
EQ-5D-5L Baseline (Day 1) utility score			
Mean (SD)	0.51 (0.23)	0.54 (0.20)	0.52 (0.21)
Median (Q1, Q3)	0.55 (0.39, 0.69)	0.58 (0.46, 0.70)	0.57 (0.42, 0.69)
Min, Max	(-0.19, 1.00)	(0.04, 0.85)	(-0.19, 1)

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

* Baseline scores calculated as the mean of Screening and Day 1 scores
BID: twice daily; BMI: Body mass index

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

Table 2. Summary of MSWS-12v1 Changes Scores* Across PGIC categories

*Method 1: Analysis included each subject's **mean** MSWS-12v1 change score from prior visit for all visits within the PGIC category*

PGIC category	N	Range	Mean (SD)	Median (Inter-Quartile Range: Q1, Q3)
1 = very much worse	4	1.0, 10.4	5.7 (4.4)	5.7 (2.1, 9.4)
2 = much worse	18	-4.2, 31.4	10.6 (10.1)	9.4 (1.6, 16.7)
3 = slightly worse	46	-22.9, 24.0	4.4 (8.7)	4.2 (0, 9.9)
4 = unchanged	120	-16.7, 22.2	0.5 (7.0)	0.0 (-2.7, 4.2)
5 = slight improvement	**81**	**-35.4, 16.7**	**-6.6 (8.2)**	**-6.3 (-12.5, -2.1)**
6 = much improved	20	-29.2, 11.5	-6.8 (12.2)	-4.8 (-13.0, 2.1)
7 = very much improved	3	-8.3, 1.0	-4.5 (4.9)	-6.3 (-8.3, 1.0)

*Method 2: Analysis included each subject's **median** MSWS-12v1 change score from prior visit for all visits within the PGIC category*

PGIC category	N	Range	Mean (SD)	Median (Inter-Quartile Range: Q1, Q3)
1 = very much worse	4	0.0, 10.4	5.5 (4.8)	5.7 (1.6, 9.4)
2 = much worse	18	-4.2, 31.4	10.5 (10.0)	9.4 (2.1, 16.7)
3 = slightly worse	46	-22.9, 24.0	4.6 (9.0)	4.2 (0.0, 11.5)
4 = unchanged	120	-18.8, 27.1	0.3 (7.3)	0.0 (-2.6, 3.7)
5 = slight improvement	**81**	**-35.4, 16.7**	**-6.4 (8.8)**	**-5.2 (-10.4, -2.1)**
6 = much improved	20	-29.2, 11.5	-6.0 (12.4)	0.0 (-13.0, 2.1)
7 = very much improved	3	-8.3, 1.0	-4.5 (4.9)	-6.3 (-8.3, 1.0)

* A single observation per subject is used within each PGIC category

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

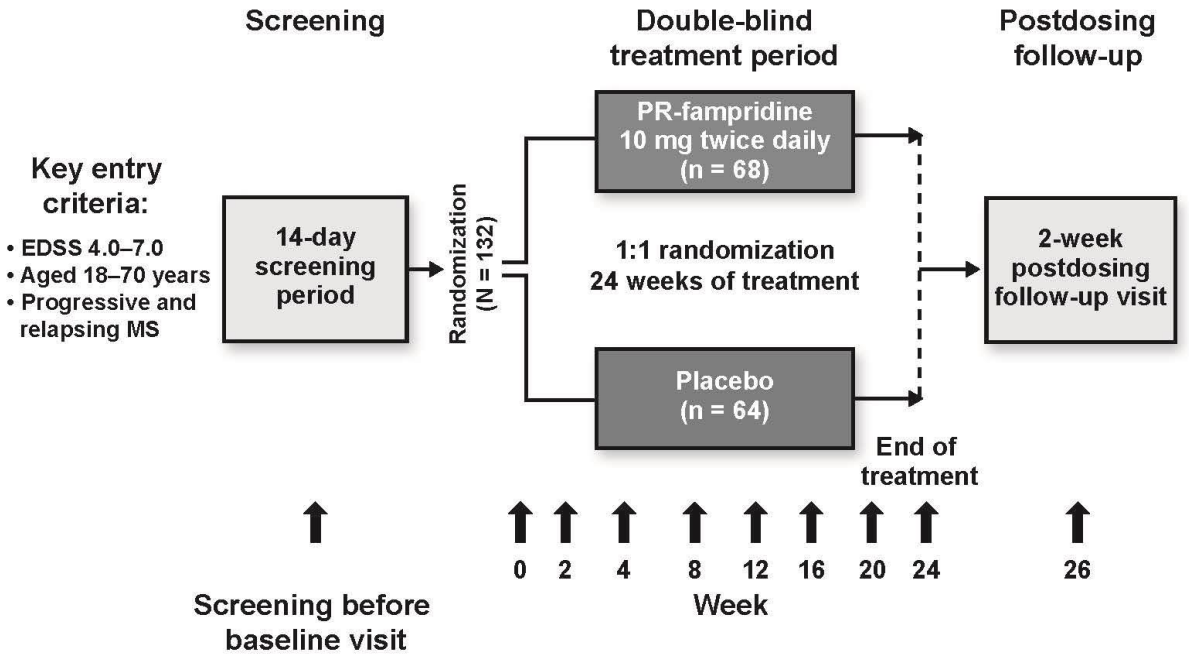
Table 3. MSWS-12v1 Responder Definition Estimates Summary

	Measures	Important Change Level (Number of subjects contributing change scores or data)	Associated MSWS-12v1 Change Estimate
Anchor-based	**PGIC:**		
	Method 1	Slightly Improved (81)	-6.6 and -6.3 (mean and median)
	Method 2		-6.4 and -5.2 (mean and median)
	EQ-5D-5L Mobility Question	1 Point Median Improvement (21)	-13.4 and -9.7 (mean and median)
Distribution-based	**Standard Error of Measurement (SEM)**	1 Standard Error of Measurement (SEM) (132)	-6.8

This is an author's draft of an accepted article submitted and published in Multiple Sclerosis Journal

DOI: http://mso.sagepub.com/content/1/2055217315596993.full

Figure 1. Study Design Schematic



EDSS: Expanded Disability Status Scale
MSWS-12v1 was administered at Screening, Day 1, Weeks 2, 4, 8, 12, 16, 20, 24, 26, and each unscheduled visit
PGIC was administered at Weeks 2, 4, 8, 12, 16, 20, and 24
EQ-5D-5L was administered at Day 1, Weeks 4, 8, 12, 16, 20, and 24
MSIS-29v1 was administered at Screening, Day 1, Weeks 2, 4, 8, 12, 16, 20, 24, and each unscheduled visit