Supplementary Material

Assessment of Clinically Meaningful Improvements in Self-reported Walking Ability in Participants with Multiple Sclerosis: Results from the Randomized, Double-blind, Phase III ENHANCE Trial of Prolonged-release Fampridine

Jeremy Hobart,[1] Tjalf Ziemssen,[2] Peter Feys,[3] Michael Linnebank,[4] Andrew D. Goodman,[5] Rachel Farrell,[6] Raymond Hupperts,[7] Andrew R. Blight,[8] Veronica Englishby,[9] Manjit McNeill,[9] Ih Chang,[10] Gabriel Lima,[10] Jacob Elkins,[10] on behalf of the ENHANCE study investigators

1 Plymouth University Peninsula Schools of Medicine and Dentistry, University Hospitals Plymouth NHS Trust, Plymouth, UK

2 Center of Clinical Neuroscience, Carl Gustav Carus University Clinic, Technical University of Dresden, Dresden, Germany

3 BIOMED-REVAL, University of Hasselt, Diepenbeek, Belgium

4 HELIOS Klinik Hagen-Ambrock, University Witten/Herdecke, Hagen, Germany

5 School of Medicine and Dentistry, University of Rochester, Rochester, NY, USA

6 National Hospital for Neurology and Neurosurgery, University College London Hospitals and University College London Institute of Neurology, London, UK

7 Zuyderland Medical Center, Sittard, the Netherlands

8 Acorda Therapeutics, Inc., Ardsley, NY, USA

9 Biogen, Maidenhead, UK

10 Biogen, Cambridge, MA, USA

Correspondence to: Dr Jeremy Hobart, Plymouth University Peninsula Schools of Medicine and Dentistry, N13 ITTC Building, Plymouth Science Park, Plymouth, Devon, PL6 8BX, UK; email: jeremy.hobart@plymouth.ac.uk; telephone: +44 1752 430055

Journal: *CNS Drugs*

Table S1 ENHANCE study investigators

Site location	Principal investigator
UMHAT 'Dr. Georgi Stranski,' Pleven, Bulgaria	Plamen Bozhinov
Military Medical Academy - MHAT, Sofia, Bulgaria	Krasimir Genov
MHAT 'SofiaMed,' Sofia, Bulgaria	Rosen Ikonomov
City Clinic UMHAC EOOD, Sofia, Bulgaria	Marko Klissurski
MHAT 'Avis Medica' OOD, Pleven, Bulgaria	Hristo Lilovski
UMHAT 'Sv. Georgi,' Plovdiv, Bulgaria	Maria Manova-Slavova
First MHAT, Sofia, Bulgaria	Dimitar Maslarov
MHATNP 'Sv.Naum,' Sofia, Bulgaria	Ivan Milanov
MHAT 'Tokuda Hospital Sofia,' Sofia, Bulgaria	Ivan Staikov
MHATNP 'Sv.Naum,' Sofia, Bulgaria	Paraskeva Stamenova
UMHAT 'Alexandrovska,' Sofia, Bulgaria	Ivaylo Tarnev
UMHAT 'Alexandrovska,' Sofia, Bulgaria	Latchezar Traykov
Military Medical Academy - MHAT, Pleven, Bulgaria	Plamen Tzvetanov
Fakultni nemocnice u sv. Anny v Brne, Brno, Czech Republic	Michal Dufek
Fakultni nemocnice Ostrava, Havirov, Czech Republic	Pavel Hradilek
VFN v Praze, Praha 2, Czech Republic	Jana Preiningerova Lizrova
Fakultni nemocnice v Motole, Praha 5, Czech Republic	Eva Houzvickova
Nemocnice Pardubickeho kraje, a.s., Pardubice, Czech Republic	Alena Novotna
Nemocnice Jihlava, Jihlava, Czech Republic	Ondrej Skoda
Krajska zdravotni, a.s. - Nemocnice Teplice, o.z., Teplice, Czech Republic	Marta Vachova
Poliklinika Chocen, a.s., Chocen, Czech Republic	Martin Valis
Fakultni nemocnice Brno, Brno, Czech Republic	Yvonne Benesova
Neuro NEO Oy, Turku, Finland	Juha-Pekka Erälinna
Oulun yliopistollinen sairaala, Oulu, Finland	Irma Keskinarkaus
Terveystalo Kamppi, Helsinki, Finland	Juha Multanen
FinnMedi Oy, Tampere, Finland	Maritta Ukkonen
IRCCS Centro Neurolesi Bonino Pulejo, Messina, Italy	Placido Bramanti
Nazionale Neurologico 'C., Milano, Italy	Paolo Confalonieri
Seconda Università degli Studi di Napoli, Napoli, Italy	Antonio Gallo
Azienda Ospedaliera Sant' Antonio Abate, Gallarate, Italy	Angelo Ghezzi
Azienda Ospedaliero Universitaria S. Luigi Gonzaga, Via Regione Gonzole, Italy	Simona Malucchi
Azienda Ospedaliero Universitaria Policlinico Paolo Giaccone, Palermo, Italy	Giuseppe Salemi
Azienda Ospedaliera Sant'Andrea - Università di Roma La Sapienza, Roma, Italy	Marco Salvetti
Vilnius University Hospital Santariskiu Clinic, Public Institution, Vilnius, Lithuania	Rasa Kizlaitiene, and Valmantas Budrys
Klaipeda University Hospital, Public Institution, Klaipeda, Lithuania	Lina Malciene
Siauliai Republican Hospital, Siauliai, Lithuania	Sigla Sceponaviciute
Hospital of Lithuanian University of Health Sciences Kaunas Clinics, Kaunas, Lithuania	Antanas Vaitkus
Orbis Medisch Centrum, Sittard-Geleen, the Netherlands	Raymond Hupperts
Amphia Ziekenhuis, Breda, the Netherlands	Ronald van Dijl
Jeroen Bosch Ziekenhuis, 's-Hertogenbosch, the Netherlands	Erik van Munster
Regionalny Szpital Specjalistyczny im. dr Wladyslawa Bieganskiego, Grudziadz, Poland	Robert Bonek
mMED Maciej Czarnecki, Warszawa, Poland	Maciej Czarnecki
Centrum Kompleksowej Rehabilitacji Sp. Z o.o., Konstancin-Jeziorna, Poland	Lidia Darda-Ledzion
Centrum Medyczne Medyk, Rzeszow, Poland	Mirosław Dziki
COPERNICUS Podmiot Leczniczy Sp. z o. o., Gdansk, Poland	Waldemar Fryze
Neurologiczny NZOZ Centrum Leczenia SM, Plewiska, Poland	Hanka Hertmanowska, and Adriana Chelminiak
Neuro-Care Site Management Organization Gabriela Klodowska-Duda, Katowice, Poland	Gabriela Klodowska-Duda
M.A. – LEK A.M.Maciejowscy SC., Katowice, Poland	Maciej Maciejowski
MCD Medical, Krakow, Poland	Ryszard Nowak
Prywatny Gabinet Neurologiczny K.Selmaj, Lodz, Poland	Krzysztof Selmaj
Wojewodzki Specjalistyczny Szpital w Olsztynie, Olsztyn, Poland	Andrzej Tutaj
Miedzyleski Szpital Specjalistyczny, Warszawa, Poland	Jacek Zaborski
Uslugi Lekarskie S.C., Warszawa, Poland	Witold Palasik
Krakowska Akademia Neurologii Sp z o.o. Centrum Neurologii Klinicznej, Krakow, Poland	Andrzej Szczudlik
NZOZ Profilaktyka Wladyslaw Pierzchala, Katowice, Poland	Krystyna Pierzchala
Uniwersytecki Szpital Kliniczny w Bialymstoku, Bialystok, Poland	Wieslaw Drozdowski
Resmedica NZOZ Kielce, Kielce, Poland	Elzbieta Jasinska
SAIH 'Kemerovo Regional Clinical Hospital,' Kemerovo, Russia	Elena Arefyeva
SBHI of Nizhny Novgorod region 'City Clinical Hospital #3,' Nizhniy Novgorod, Russia	Anna Belova
SBHI 'CCH #24 of Department for Healthcare of Moscow,' Moscow, Russia	Alexey Boyko
Saint-Petersburg SHI 'City Clinical Hospital #31,' Saint-Petersburg, Russia	Evgeny Evdoshenko

Republican Clinical Hospital for Rehabilitation of Healthcare Ministry of Republic of Tatarstan, Kazan, Russia	Farit Khabirov
FSI Siberian District Medical Centre of Federal agency, Krasnoyarsk, Russia	Dmitry Pokhabov
Regional Multiple Sclerosis Centre b/o CC ECM «Neftyanik» 12, Tyumen, Russia	Stella Sivertseva
NIH 'Central Clinical Hospital # 2' n.a.N.A. Semashko of OJSC 'Russian Railways,' Moscow, Russia	Olga Vorobyova
Clinical Center Kragujevac, Kragujevac, Serbia	Tatjana Boskovic Matic
Clinical Center of Serbia, Belgrade, Serbia	Jelena Drulovic
Military Medical Academy, Belgrade, Serbia	Dragana Obradovic
Clinical Center Nis, Nis, Serbia	Slobodan Vojinovic
Universitaetsspital Basel, Neurologische Klinik und Poliklinik, Basel, Switzerland	Tobias Derfuss
St Peter's Hospital, Chertsey, UK	David Barnes
The National Hospital for Neurology & Neurosurgery, London, UK	Jeremy Chataway
Queen's Hospital, Romford, UK	Abhijit Chaudhuri
Queen's Medical Centre, Nottingham, UK	Cris Constantinescu
Royal Devon and Exeter Hospital (Wonford), Exeter, UK	Timothy Harrower
Derriford Hospital, Plymouth, UK	Jeremy Hobart
Norfolk and Norwich University Hospital, Norwich, UK	Martin Lee
Morriston Hospital, Swansea, UK	Owen Pearson
Salford Royal, Salford, UK	David Rog
Royal London Hospital, London, UK	Benjamin Turner
Queen Elizabeth University Hospital Campus, Glasgow, UK	Stewart Webb
Royal Free Hospital, London, UK	Heather Wilson
Queen Elizabeth Hospital, Birmingham, UK	John Woolmore
Associates in Neurology, P.S.C., Lexington, KY, USA	Gregory Anderson
Barrow Neurological Institute, Phoenix, AZ, USA	Aimee Borazanci
Ohio State University Medical Center, Columbus, OH, USA	Aaron Boster, Jacqueline Nicholas, and Michael Racke
Minneapolis Clinic of Neurology, Golden Valley, MN, USA	Jonathan Calkwood
Novex Clinical Research, LLC, New Bedford, MA, USA	Mushtaque Chachar
Carolinas Healthcare System, Charlotte, NC, USA	Benjamin Brooks, Jill Conway, and Donna Graves
University of Rochester, Rochester, NY, USA	Andrew Goodman and Lawrence Samkoff
Mercy Clinic Neurology , Town and Country Commons, Chesterfield, MO, USA	Barbara Green
Meridien Research, Tampa, FL, USA	Cynthia Huffman
Wayne State University, Detroit, MI, USA	Omar Khan
North Central Neurology Associates, P.C., Cullman, AL, USA	Christopher Laganke
Bradenton Research Center, Inc., Bradenton, FL, USA	William McElveen
Alabama Neurology Associates, Birmingham, AL, USA	Emily Riser
Metrolina Neurological Assoc., PA, Charlotte, NC, USA	Allan Ryder-Cook
Pacific Research Network, Inc, San Diego, CA, USA	Stephen Thein
Premiere Research Institute, West Palm Beach, FL, USA	Paul Winner
Blue Ridge Research Center, LLC, Roanoke, VA, USA	John Burch
Compass Research, LLC, Orlando, FL, USA	Ira Goodman
University of South Florida Rothman Center of Neuropsychiatry, Tampa, FL, USA	Derrick Robertson
Infinity Clinical Research, LLC, Sunrise, FL, USA	Barry Cutler

Supplementary Methods

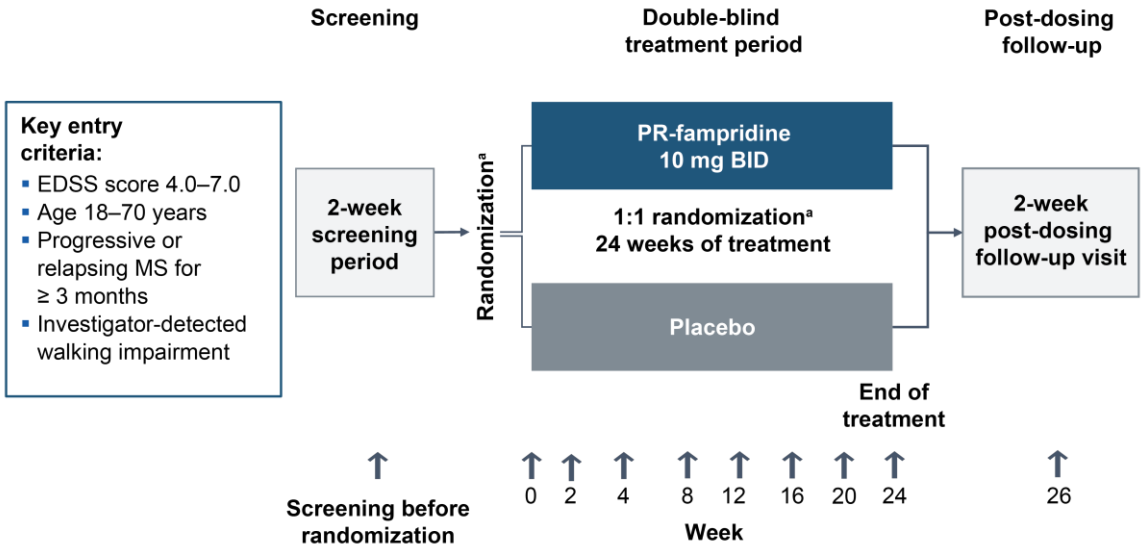

Fig. S1 Study design. [a]Participants were randomized using an interactive voice/web response system and were stratified by EDSS score at screening (≤ 6.0 or > 6.0) and prior aminopyridine use; caps were applied so that the proportions of participants with EDSS score > 6.0 and prior aminopyridine use did not exceed ~ 35% and ~ 10%, respectively. Arrows indicate study visits. The first participant received treatment on September 29, 2014 and the last patient's last study visit was February 11, 2016. *BID* twice daily, *EDSS* Expanded Disability Status Scale, *MS* multiple sclerosis, *PR* prolonged-release.

Inclusion and Exclusion Criteria

To be eligible to participate in this study, candidates were required to meet the following eligibility criteria at the screening visit or at the time point specified in the individual eligibility criterion listed (potential participants who failed screening could be re-screened once):

1. Ability to understand the purpose and risks of the study and provide signed and dated informed consent and authorization to use protected health information in accordance with national and local patient privacy regulations.
2. Age 18–70 years, inclusive, at the time of informed consent.
3. Female participants of childbearing potential were required to have a negative urine pregnancy test at the screening visit and on day 1. All participants were required to agree to practice effective contraception during the study and to be willing and able to continue contraception for 30 days after their last dose of study treatment. Effective contraception methods were defined in the protocol.
4. Had a diagnosis of primary progressive multiple sclerosis (MS), secondary progressive MS, progressive-relapsing MS, or relapsing-remitting MS, per revised McDonald Committee criteria [1, 2] as defined by Lublin and Reingold [3], of ≥ 3 months' duration.
5. Had an Expanded Disability Status Scale score 4.0–7.0, inclusive.
6. Had walking impairment, as deemed by the investigator.
7. Ability to understand and comply with the requirements of the protocol.

Candidates were excluded from study entry if any of the following exclusion criteria existed at the screening visit or at the time point specified in the individual criterion listed:

1. History of human immunodeficiency virus.
2. Presence of acute or chronic hepatitis. Participants who had evidence of prior hepatitis infection that was serologically confirmed as resolved were not excluded from study participation.
3. Known allergy to fampridine, pyridine-containing substances, or any of the inactive ingredients in the prolonged-release (PR) fampridine tablet.
4. Any history of seizure, epilepsy, or other convulsive disorder, with the exception of febrile seizures in childhood.
5. Creatinine clearance < 80 mL/min.
6. History of malignant disease, including solid tumors and hematologic malignancies (with the exception of basal cell and squamous cell carcinomas of the skin that were completely excised and considered cured) within the 5 years before the screening visit or at any time during the screening period.
7. Onset of MS exacerbation within 60 days before the screening visit, or at any time during the screening period.
8. History of any major surgical intervention (with the exception of skin biopsy) within the 30 days before the screening visit or day 1, or at any time during the screening period.
9. Any non-MS–related condition or factor (as determined by the investigator) that was likely to interfere with walking ability, including, but not limited to, previous major surgery of the foot, leg, or hip; any significant trauma; or known peripheral neuropathy of the lower limb.
10. Presence of pulmonary disease, including, but not limited to, chronic obstructive pulmonary disease that could impede the participant's daily activities (as determined by the investigator).
11. Presence of any psychiatric disorder, including clinical depression, that was likely to interfere with the participant's participation in the study (as determined by the investigator).
12. Uncontrolled hypertension (as determined by the investigator) at the screening visit or at any time during the screening period.
13. History of any clinically significant cardiac, endocrinologic, hematologic, immunologic, metabolic, urologic, neurologic (except for MS, but including events indicative of a potentially lower seizure threshold), dermatologic, or other major disease (as determined by the investigator).
14. Clinically significant abnormal laboratory values (as determined by the investigator).
15. Body mass index ≥ 40 kg/m^2.
16. History of severe allergic or anaphylactic reactions.
17. Use of off-label MS treatment including rituximab, daclizumab, or antibody (except natalizumab) within the 3 months before the screening visit, at any time during the screening period, or scheduled for use during study participation.
18. Use of mitoxantrone or cyclophosphamide within the 3 months before the screening visit, at any time during the screening period, or scheduled for use during study participation.
19. Initiation of natalizumab or alemtuzumab treatment, or any change in the participant's dose or regimen of natalizumab or alemtuzumab, within the 3 months before the screening visit, or at any time during the screening period.

20. Initiation of treatment with, or any change in the participant's dose or regimen of, interferon beta-1b, interferon beta-1a, fingolimod, teriflunomide, glatiramer acetate, or dimethyl fumarate within the 30 days before the screening visit or at any time during the screening period.
21. Pulsed steroid treatment within the 60 days before the screening visit or at any time during the screening period.
22. Any change in the participant's medication dose or regimen for the treatment of fatigue or depression within the 30 days before the screening visit or at any time during the screening period.
23. Any change in prophylactic treatment for pain with antidepressants or anticonvulsants prescribed for this purpose within 30 days before the screening visit or at any time during the screening period.
24. Any change in the participant's dose or regimen of antispastic agents within the 7 days before the screening visit or at any time during the screening period.
25. Treatment with an investigational drug within the 30 days (or seven half-lives, whichever is longer) before the screening visit or at any time during the screening period.
26. Treatment with any aminopyridine (fampridine, 4-aminopyridine, or 3,4-diaminopyridine in any formulation) within the 30 days before the screening visit or at any time during the screening period.
27. Treatment with organic cation transporter 2 inhibitors (list provided in the Study Reference Manual) within five half-lives before the screening visit or at any time during the screening period.
28. History of drug or alcohol abuse (as defined by the investigator) within the 2 years before the screening visit.
29. Female participants who were currently pregnant or who were considering becoming pregnant while participating in the study.
30. Female participants who were currently breastfeeding.
31. Inability to comply with study requirements.
32. Participants who planned to participate in another clinical study (including any observational studies) during the course of the current study.
33. Other unspecified reasons that, in the opinion of the investigator or Biogen, made the participant unsuitable for enrollment.

Supplementary Methods (continued)

Pre-treatment, on-treatment, and mean change distributional statistics were reported for each clinical outcome assessment (COA). Two effect size calculations, standardized change scores converting raw change scores into standard deviation (SD) units, were also calculated: Cohen's effect size (mean change / pre-treatment SD and mean change / pooled SD) and standardized response means (SRMs; mean change / SD change). By standardizing raw change scores, effect sizes enable head-to-head comparisons of different COAs with different numbers of items, item response categories, ranges, variances, and units. By convention, effect sizes are interpreted using Cohen's criteria: 0.20 is the threshold for a clinically small change; 0.50 the threshold for a clinically moderate change; and 0.80 the threshold for a clinically large change [4].

Table S3 enables three important comparisons: PR-fampridine 12-item Multiple Sclerosis Walking Scale (MSWS-12) responders (≥ 8-point mean improvement from baseline over 24 weeks) vs. the total placebo group; PR-fampridine MSWS-12 responders vs. the total PR-fampridine group; and PR-fampridine MSWS-12 non-responders vs. placebo. The comparison of PR-fampridine MSWS-12 non-responders vs. placebo is particularly important, as it indicates the similarity in outcomes between PR-fampridine MSWS-12 non-responders and placebo-treated participants. Theoretically, a MSWS-12 non-responder will not have a physiological response to PR-fampridine; therefore, a PR-fampridine MSWS-12 non-responder is the equivalent of a placebo-treated participant. As such, the similarities in outcomes of these two groups indicates the extent to which the MSWS-12 responder definition distinguishes 'true' responders from 'true' non-responders.

Supplementary Results

MSWS-12 results (Table S3) show that the sample-to-scale targeting was reasonable because the pre-treatment means (PR-fampridine, 63.6; placebo, 65.4) were relatively near the midpoint score of the MSWS-12 (50 points), and the floor/ceiling effects were minimal. However, the distribution of MSWS-12 scores were skewed towards higher scores (worse walking ability). The mean change from baseline and effect sizes in the PR-fampridine and placebo groups indicated an improvement in walking ability during the on-treatment period; the magnitude of improvements were twice as large for PR-fampridine–treated than placebo-treated participants. Both PR-fampridine group effect sizes comfortably exceeded the threshold for a clinically small improvement in walking (0.20), and the SRM implied a near clinically moderate benefit on walking ability. The improvements in walking ability in the placebo group observed over 24 weeks did not meet Cohen's effect size threshold of clinically small, and just reached the threshold of clinically small for the SRM.

PR-fampridine MSWS-12 responders showed a mean improvement of –20.4 points from baseline in MSWS-12 score and PR-fampridine responder/non-responders had a mean change difference of –23.7 points; essentially 25% of the entire scale range and a magnitude of improvement 2.5 times as large as the clinically meaningful 8-point threshold [5]. The PR-fampridine MSWS-12 responder effect size (Cohen's effect size = pre-treatment SD –1.0, pooled SD –1.9; SRM = –1.7) were very large. Both the Cohen and SRM effect sizes for PR-fampridine MSWS-12 responders notably exceeded the threshold for a clinically large walking improvement (0.80) vs. PR-fampridine MSWS-12 non-responders. Naturally, MSWS-12 change scores and effect sizes in MSWS-12 non-responders were expected to be small; by definition, the maximum possible change score is –7.9 points.

Timed Up and Go (TUG) time (seconds) measured the time taken to stand up, walk 3 meters, and return to the seat (Table S3); therefore, the ceiling/floor effects were not applicable. Improvements were observed in both the PR-fampridine and placebo groups over 24 weeks, and the magnitude of change observed in PR-fampridine was twice that of placebo. However, changes in TUG time in the PR-fampridine, placebo, PR-fampridine MSWS-12 responder, and PR-fampridine MSWS-12 non-responder groups over 24 weeks did not meet Cohen's effect size threshold for clinical change, whereas the SRM indicated clinically small changes in TUG time for the PR-fampridine–treated participants (–0.28) and PR-fampridine MSWS-12 responders (–0.44).

Results for the Multiple Sclerosis Impact Scale 20-item physical impact subscale (MSIS-29 PHYS) show that the sample-to-scale-targeting was adequate as all observed pre-treatment mean scores were at the scale midpoint (50 points) and floor/ceiling effects were minimal (Supplementary table S3). Both the PR-fampridine and placebo groups recorded improvements in physical functioning during treatment from baseline. PR-fampridine–treated participants had a greater mean change scores (ratio 1.56) from baseline and larger effect sizes (ratios 2.00; 1.50) than placebo-treated participants; both effect sizes comfortably exceeded clinically small improvements in physical functioning and were near moderate (Cohen's effect size) and moderate (SRM). However, the mean change score from baseline and effect size differences between the PR-fampridine and placebo groups were less notable.

In the MSWS-12 responder analyses, PR-fampridine MSWS-12 responders showed large mean MSIS-29 PHYS score changes from baseline (16% of scale range) and effect sizes, implying clinically large improvements in physical function. PR-fampridine MSWS-12 non-responders had near zero MSIS-29 PHYS mean change scores and effect size.

Table S3 reports the PR-fampridine vs. placebo scores for the 14-item Berg Balance Scale (BBS) over 24 weeks. Sample-to-scale targeting was suboptimal; the observed pre-treatment mean BBS scores (PR-fampridine, 40.6; placebo, 40.2) were above the scale midpoint (28 points) and skewed notably towards better balance. However, the floor/ceiling effects of the BBS were negligible. The modified intention-to-treat (ITT) comparison of BBS showed a numerical increase in balance in both the PR-fampridine and placebo groups from baseline, but the magnitudes of change of the mean scores reported by Cohen's effect size were small. The SRM implied that the magnitude of improved balance with PR-fampridine was moderately large over 24 weeks, and in non-responders was small.

In the MSWS-12 responder analysis, the magnitudes of change from baseline and difference in effect sizes of BBS scores in PR-fampridine MSWS-12 responders were twice as large as that for non-responders. The BBS effect size scores in PR-fampridine MSWS-12 responders implied that improvements in balance exceeded the criteria for clinically small (Cohen's effect size) and clinically moderate (SRM). The effect sizes for PR-fampridine MSWS-12 non-responders implied that improvements in balance were negligible (Cohen's effect size) or small to moderate (SRM).

Table S3 shows the PR-fampridine and placebo scores for the 56-item version of ABILHAND. Sample-to-scale targeting was poor; pre-treatment mean ABILHAND scores (PR-fampridine, 86.9; placebo, 84.3) were above the ABILHAND midpoint (50 points) and very skewed towards greater hand functioning. The mean on-treatment score change from baseline and the Cohen's and SRM effect sizes showed very small and similar numerical improvements in manual ability in both the PR-fampridine and placebo groups.

In the MSWS-12 responder analysis, the magnitude of mean ABILHAND score change from baseline of the PR-fampridine MSWS-12 responders were ~ 10 times larger than for PR-fampridine MSWS-12 non-responders. The PR-fampridine MSWS-12 responder effect sizes exceeded the criteria for clinically small improvements in manual ability; however, the MSWS-12 non-responder effects sizes were near zero, implying no improvement in manual ability.

Table S2 Concomitant medication and non-drug therapy during the study

Concomitant therapy	PR-fampridine (*n* = 316)	Placebo (*n* = 319)
Any concomitant medication	276 (87)	287 (90)
Most common concomitant medications[a]		
Baclofen	65 (21)	65 (20)
Colecalciferol	47 (15)	48 (15)
Tizanidine	36 (11)	37 (12)
Ibuprofen	33 (10)	31 (10)
Methylprednisolone	35 (11)	29 (9)
Paracetamol	31 (10)	30 (9)
Any concomitant non-drug therapy[b]	43 (14)	51 (16)
Most common concomitant non-drug therapies[b]		
Physiotherapy	16 (5)	19 (6)
Bladder catheterization	0	9 (3)
Rehabilitation therapy	3 (< 1)	5 (2)

PR prolonged-release.

[a]Medication used in ≥ 10% of participants in either group.

[b]Therapy received in ≥ 2% of participants in either group.

Table S3 Clinical outcome assessments in the modified intention-to-treat sample and by PR-fampridine MSWS-12 responder (≥ 8-point mean improvement) status

	Modified intention-to-treat comparison									
	MSWS-12		ABILHAND		MSIS-29 PHYS		BBS		TUG time s	
	PR-fampridine (*n* = 315)	Placebo (*n* = 318)	PR-fampridine (*n* = 312)	Placebo (*n* = 315)	PR-fampridine (*n* = 315)	Placebo (*n* = 318)	PR-fampridine (*n* = 315)	Placebo (*n* = 318)	PR-fampridine (*n* = 315)	Placebo (*n* = 318)
Pre-treatment score[a]										
Mean (SD)	63.6 (21.7)	65.4 (21.9)	86.9 (15.8)	84.3 (16.5)	52.4 (21.1)	55.3 (21.0)	40.6 (11.6)	40.2 (11.8)	24.9 (26.6)	27.1 (42.03)
Range	0.0 to 100.0	0.0 to 100.0	0.9 to 100.0	26.0 to 100.0	0.0 to 98.3	3.3 to 95.8	6.0 to 56.0	4.0 to 56.0	6.3 to 239.8	0.0 to 436.8
Floor/ceiling effect, *n* (%)	4 (1) / 1 (< 1)	5 (2) / 1(< 1)	0 / 60 (19)	0 / 59 (19)	0 / 1 (< 1)	0 / 0	0 / 3 (< 1)	0 / 2 (< 1)		
On-treatment score										
Mean (SD)	56.7 (24.5)	62.0 (23.4)	88.6 (13.9)	86.0 (15.5)	45.0 (22.2)	50.5 (22.2)	42.2 (11.6)	41.5 (12.1)	22.5 (26.6)	26.0 (40.1)
Range	0.9 to 100.0	3.8 to 100.0	25.5 to 100.0	30.3 to 100.0	0.2 to 98.1	1.7 to 96.0	6.0 to 56.0	2.2 to 56.0	4.6 to 270.4	0.0 to 403.1
Floor/ceiling effect, *n* (%)[b]	3 (< 1) / 0	6 (2) / 0	0 / 54 (17)	0/ 50 (16)	0 / 0	0 / 0	0 / 2 (< 1)	0 / 4 (1)		
Change of ≥ 8 points over 24 weeks										
Participants with improvement, *n* (%)[c]	136 (43.2)	107 (33.6)								
p value vs. placebo[d]	*p* = 0.006									
Odds ratio vs. placebo (95% CI)[d]	1.61 (1.15 to 2.26)									
Risk difference (95% CI)[d]	0.104 (0.03 to 0.18)									
Change from baseline over 24 weeks[e]										
Mean (SD)	–7.0 (15.8)	–3.4 (15.3)	1.7 (8.2)	1.6 (8.4)	–7.5 (13.6)	–4.8 (13.2)	1.7 (3.3)	1.3 (3.8)	–2.4 (8.6)	–1.2 (12.7)
Range	–68.1 to 70.8	–60.6 to 85.9	–26.5 to 80.2	–40.1 to 46.3	–53.3 to 34.1	–62.3 to 48.8	–13.4 to 16.5	–20.5 to 18.8	–79.9 to 33.5	–89.7 to 82.8
Floor/ceiling effect, *n* (%)[f]	1 (< 1) / 0	3 (< 1) / 0	0 / 41 (13)	0 / 40 (13)	0 / 0	0 / 0	0 / 2 (< 1)	0 / 1 (< 1)		
Effect sizes										
Cohen's effect size (mean / SD)[g]	–0.32 (–6.9 / 21.7)	–0.15 (–3.4 / 21.9)	0.11 (1.7 / 15.8)	0.10 (1.6 / 16.5)	–0.35 (–7.5 / 21.1)	–0.23 (–4.8 / 21.0)	0.14 (1.7 / 11.6)	0.11 (1.3 / 11.8)	–0.09 (–2.4 / 26.6)	–0.03 (–1.2 / 42.03)
Cohen's effect size (mean / SD)[h]	–0.44 (–6.9 / 15.6)	–0.22 (–3.4 / 15.6)	0.20 (1.7 / 8.3)	0.19 (1.6 / 8.3)	–0.56 (–7.5 / 13.4)	–0.36 (–4.8 / 13.4)	0.48 (1.7 / 3.6)	0.37 (1.3 / 3.6)	–0.22 (–2.4 / 10.9)	–0.11 (–1.2 / 10.9)
Standardized response mean (mean / SD)[i]	–0.44 (–6.9 / 15.8)	–0.22 (–3.4 / 15.3)	0.21 (1.7 / 8.2)	0.19 (1.6 / 8.4)	–0.55 (–7.5 / 13.6)	–0.36 (–4.8 / 13.2)	0.51 (1.7 / 3.3)	0.33 (1.3 / 3.8)	–0.28 (–2.4 / 8.6)	–0.09 (–1.2 / 12.7)

	MSWS-12 responder (≥8-point improvement) analysis									
	MSWS-12		ABILHAND		MSIS-29 PHYS		BBS		TUG time s	
	PR-fampridine responders (*n* = 136)	PR-fampridine non-responders (*n* = 179)	PR-fampridine responders (*n* = 136)	PR-fampridine non-responders (*n* = 179)	PR-fampridine responders (*n* = 136)	PR-fampridine non-responders (*n* = 179)	PR-fampridine responders (*n* = 136)	PR-fampridine non-responders (*n* = 179)	PR-fampridine responders (*n* = 136)	PR-fampridine non-responders (*n* = 179)
Pre-treatment score[a]										
Mean (SD)	64.2 (20.1)	63.2 (22.8)	86.6 (17.0)	87.2 (14.9)	50.8 (21.0)	53.7 (21.2)	41.5 (10.7)	39.8 (12.3)	21.8 (19.8)	27.3 (30.7)
Range	12.4 to 97.9	0.0 to 100.0	0.9 to 100.0	14.3 to 100.0	0.0 to 98.3	6.7 to 95.8	9.0 to 56.0	6.0 to 56.0	6.8 to 112.3	6.3 to 239.8
Floor/ceiling effect, *n* (%)	0 / 0	4 (2) / 1 (< 1)	0 / 27 (20)	0 / 33 (18)	0 / 1 (< 1)	0 / 0	0 / 2 (1)	0 / 1 (< 1)		

MSWS-12 responder (≥8-point improvement) analysis (cont.)

	MSWS-12		ABILHAND		MSIS-29 PHYS		BBS		TUG time s	
	PR-fampridine responders (*n* = 136)	PR-fampridine non-responders (*n* = 179)	PR-fampridine responders (*n* = 136)	PR-fampridine non-responders (*n* = 179)	PR-fampridine responders (*n* = 136)	PR-fampridine non-responders (*n* = 179)	PR-fampridine responders (*n* = 136)	PR-fampridine non-responders (*n* = 179)	PR-fampridine responders (*n* = 136)	PR-fampridine non-responders (*n* = 179)
On-treatment score										
Mean (SD)	43.8 (21.1)	66.5 (22.2)	90.2 (13.1)	87.5 (14.3)	34.7 (20.1)	52.8 (20.5)	44.0 (10.3)	40.9 (12.4)	18.2 (16.4)	25.9 (31.9)
Range	0.9 to 87.5	4.4 to 100.0	43.5 to 100.0	25.5 to 100.0	0.2 to 81.9	1.7 to 98.1	13.0 to 56.0	6.0 to 56.0	5.1 to 140.6	5.5 to 270.4
Floor/ceiling effect, *n* (%)[b]	0 / 0	3 (2) / 0	0 / 28 (21)	0 / 26 (15)	0 / 0	0 / 0	0 / 1 (< 1)	0 / 1 (< 1)		
Change from baseline over 24 weeks[e]										
Mean (SD)	–20.4 (12.1)	3.3 (9.1)	3.6 (9.5)	0.3 (6.7)	–16.1 (11.9)	–1.0 (10.9)	2.4 (3.1)	1.1 (3.3)	–3.7 (8.4)	–1.4 (8.7)
Range	–68.1 to –8.0	–7.7 to 70.8	–26.2 to 80.2	–26.5 to 41.0	–53.3 to 7.8	–37.6 to 34.1	–3.4 to 13.8	–13.4 to 16.5	–54.0 to 29.6	–79.9 to 33.5
Floor/ceiling effect, *n* (%)[f]	0 / 0	1 (< 1) / 0	0 / 22 (16)	0 / 19 (11)	0 / 0	0 / 0	0 / 1 (< 1)	0 / 1 (< 1)		
Effect size										
Cohen's effect size (mean / SD)[g]	–1.01 (–20.4 / 20.1)	0.14 (3.3 / 22.8)	0.21 (3.6 / 17.0)	0.02 (0.3 / 14.9)	–0.77 (–16.1 / 21.0)	0.05 (–1.0 / 21.2)	0.23 (2.4 / 10.7)	0.09 (1.1 / 12.3)	–0.19 (–3.7 / 19.8)	–0.05 (–1.4 / 30.7)
Cohen's effect size (mean / SD)[h]	–1.94 (–20.4 / 10.5)	0.31 (3.3 / 10.5)	0.45 (3.6 / 8.0)	0.04 (0.3 / 8.0)	–1.42 (–16.1 / 11.3)	–0.09 (–1.0 / 11.3)	0.75 (2.4 / 3.2)	0.34 (1.1 / 3.2)	–0.43 (–3.7 / 8.6)	–0.16 (–1.4 / 8.6)
Standardized response mean (mean / SD)[i]	–1.68 (–20.4 / 12.1)	0.36 (3.3 / 9.1)	0.38 (3.6 / 9.5)	0.04 (0.3 / 6.7)	–1.35 (–16.1 / 11.9)	–0.09 (–1.0 / 10.9)	0.78 (2.4 / 3.1)	0.33 (1.1 / 3.3)	–0.44 (–3.7 / 8.4)	–0.16 (–1.4 / 8.7)

Lower MSWS-12 scores indicate greater walking ability: floor score = 100, ceiling score = 0. ABILHAND score ranged from 0–100; higher scores indicate greater ability: floor score = 0; ceiling score = 100. MSIS-29 PHYS score ranged from 0–100; lower scores indicate greater ability: floor score = 100; ceiling score = 0. BBS score ranged from 0–56; higher scores indicate greater ability: floor score = 0; ceiling score = 56. TUG time was measured in s; a negative change indicates improvement from baseline (no floor or ceiling scores). Missing data were imputed using multiple imputation.

BBS Berg Balance Scale, *CI* confidence interval, *MSIS-29* Multiple Sclerosis Impact Scale, *MSWS-12* 12-item Multiple Sclerosis Walking Scale, *PHYS* physical impact subscale, *PR* prolonged-release, *SD* standard deviation, *TUG* Timed Up and Go.

[a]Pre-treatment scores calculated as the mean of the screening and baseline visits.

[b]Calculated as the number of participants with maximum/minimum scores during the mean on-treatment period.

[c]Estimated proportions based on binomial proportions.

[d]Calculated using an adjusted logistic regression model.

[e]Based on mean on-treatment values during the treatment period; calculated by subtracting the pre-treatment scores from the on-treatment scores.

[f]Calculated as the number of participants with maximum/minimum scores at both pre-treatment and mean on-treatment visits.

[g]Cohen's effect size calculated from the mean change from baseline scores divided by the pre-treatment SD.

[h]Cohen's effect size calculated from the mean change from baseline scores divided by the pooled SD.

[i]Standardized response mean equals the mean change from baseline divided by the SD change from baseline.

Table S4 Mobility outcome measures, with stratification of the PR-fampridine and placebo groups by MSWS-12 response (≥ 8-point mean improvement)

Endpoint	PR-fampridine responders (*n* = 136)	PR-fampridine non-responders (*n* = 179)	Placebo responders (*n* = 107)	Placebo non-responders (*n* = 211)
MSWS-12 score change from baseline[a]				
LSM (SE) change from baseline over 24 weeks	–20.78 (0.97)	2.29 (0.84)	–18.34 (1.05)	3.58 (0.83)
LSM difference (95% CI) vs. placebo responders	–2.45 (–4.79 to –0.11)			
LSM difference (95% CI) vs. placebo non-responders		–1.29 (–3.12 to 0.54)		
Clinically meaningful improvement (≥ 15%) in TUG speed[b]				
Participantswith improvement, %	52.4	36.6	49.5	27.2
Odds ratio (95% CI) vs. placebo responders	1.12 (0.65 to 1.93)			
Odds ratio (95% CI) vs. placebo non-responders		1.54 (0.96 to 2.45)		
TUG percentage speed change from baseline[a]				
LSM (SE) change from baseline over 24 weeks	23.95 (2.35)	10.82 (2.06)	21.68 (2.60)	7.63 (2.14)
LSM difference (95% CI) vs. placebo responders	2.27 (–3.44 to 7.97)			
LSM difference (95% CI) vs. placebo non-responders		3.19 (–1.62 to 8.01)		
MSIS-29 PHYS score[a]				
LSM (SE) change from baseline over 24 weeks	–17.48 (1.01)	–1.92 (0.87)	–14.86 (1.10)	–0.62 (0.87)
LSM difference (95% CI) vs. placebo responders	–2.62 (–5.07 to –0.17)			
LSM difference (95% CI) vs. placebo non-responders		–1.29 (–3.22 to 0.64)		
BBS score[a]				
LSM (SE) change from baseline over 24 weeks	2.58 (0.36)	1.22 (0.31)	2.38 (0.39)	0.91 (0.31)
LSM difference (95% CI) vs. placebo responders	0.21 (–0.65 to 1.07)			
LSM difference (95% CI) vs. placebo non-responders		0.31 (–0.38 to 0.99)		
ABILHAND score[a]	*n* = 133	*n* = 179	*n* = 107	*n* = 208
LSM (SE) change from baseline over 24 weeks	3.35 (0.75)	0.35 (0.65)	2.64 (0.81)	0.01 (0.65)
LSM difference (95% CI) vs. placebo responders	0.70 (–1.08 to 2.49)			
LSM difference (95% CI) vs. placebo non-responders		0.34 (–1.08 to 1.75)		

LSM (SE), LSM difference, and 95% CI vary slightly from data in Table 5 of the main manuscript because a different analysis model was fitted with different treatment groups (PR-fampridine MSWS-12 responders, PR-fampridine MSWS-12 non-responders, and placebo) compared with the model used here (PR-fampridine MSWS-12 responders, PR-fampridine MSWS-12 non-responders, placebo MSWS-12 responders, and placebo MSWS-12 non-responders); thus, the difference in random variation results in slightly different estimates.
BBS Berg Balance Scale, *CI* confidence interval, *LSM* least squares mean, *MSIS-29* Multiple Sclerosis Impact Scale, *MSWS-12* 12-item Multiple Sclerosis Walking Scale, *PHYS* physical impact subscale, *PR* prolonged-release, *SE* standard error, *TUG* Timed Up and Go.
[a]LSM, LSM difference, SE, and 95% CI calculated using a mixed model for repeated measures (missing data imputed using multiple imputation).
[b]Estimated proportion, odds ratio, and 95% CI calculated using an adjusted logistic regression model (missing data imputed using multiple imputation).

Supplementary Discussion

The ITT comparison of PR-fampridine vs. placebo implied a consistent benefit across all COAs; benefits were consistently larger for participantsrandomized to the PR-fampridine group than placebo, except for ABILHAND, where the changes from baseline in the PR-fampridine and placebo groups were similar. The magnitude of improvement from baseline in the PR-fampridine population was variable; it was largest for MSWS-12 and MSIS-29 PHYS scores, with smaller not statistically significant, but numerically greater improvements observed in BBS and ABILHAND scores. The percentage of MSWS-12 responders was 43.2% of the overall PR-fampridine population; the inclusion of scores from PR-fampridine MSWS-12 non-responders may explain why the relative differences in the magnitude of benefit between PR-fampridine and placebo participantswere small vs. scores from the PR-fampridine MSWS-12 responder population only. ITT analyses are advantageous because the randomized groups are compared; however, the comparison of PR-fampridine MSWS-12 responders vs. non-responders is important and meaningful.

The MSWS-12 responder-based analyses showed that PR-fampridine MSWS-12 responders gained clinically meaningful improvements in self-reported walking ability, physical function and manual ability, and clinician-reported dynamic balance compared with placebo, although statistical significance was not evaluated in subgroup analyses. Again, the magnitude of improvements from baseline in PR-fampridine MSWS-12 responders vs. placebo varied across the COAs. For example, SRMs implied improvements above and beyond placebo that comfortably exceeded clinically large for self-reported walking ability and physical functioning, clinically moderate for clinician-rated dynamic balance, and comfortably exceeded clinically small for manual function.

The ≥ 8-point MSWS-12 threshold of improvement used to categorize PR-fampridine MSWS-12 responders and non-responders would suggest that we expect notable differences between these groups. However, the magnitude of MSWS-12 responder change is striking; the least squares mean improvement in PR-fampridine MSWS-12 responders was –20.4 points from baseline, a 20% change in the whole scale range. Cohen's effect size and SRM of the MSWS-12 change from baseline in MSWS-12 responders exceeded unity, the SD of the distribution, and indicated improvements in walking ability that comfortably exceeded the criteria as clinically large. There was a net worsening in walking ability for MSWS-12 non-responders; therefore, the differences between the MSWS-12 responder and MSWS-12 non-responder populations were even greater.

Results from the other COAs were very encouraging, as non-walking functions measured aspects of functioning that were independent of the criterion used to define a PR-fampridine MSWS-12 responder. The improvements in physical functioning in MSWS-12 responders were particularly notable; differences in effect sizes indicated clinically large benefits, but there were also clinically meaningful benefits on dynamic balance and manual ability. Does this infer that PR-fampridine has a differential effect on different outcomes? We do not think so. A closer look at the data in Table S3 shows that targeting for the COAs differs. This is particularly important for the MSWS-12, MSIS-29, BBS, and ABILHAND scales because they have limited ranges that, by definition, restrict the potential for measuring change. In contrast, there is no upper limit for the TUG as it is a timed test.

A careful look at the pre-treatment COA score distributions and the relationship between the observed mean scores, possible scale range, and scale midpoints show important implications for interpreting findings. The examination of sample-to-scale-targeting goes beyond the simple examination of percentage floor and ceiling effects, which are valuable but only indicate participants at the absolute scale extremes and can therefore be misleading. The ability of a scale to convert true change to a change in scale score varies across the scale range: best at the center, and increasingly worse as you move away from the scale midpoint.

PR-fampridine–treated participantsshowed reductions in TUG time (seconds) that were twice that of the placebo population, with improvements in TUG time also observed in PR-fampridine MSWS-12 responders vs. non-responders. However, these differences did not translate into clinically meaningful changes as per Cohen's effect size, with clinically small SRM changes in the PR-fampridine and PR-fampridine MSWS-12 responders. The effect size results also do not agree with the secondary endpoint that demonstrated a significantly higher percentage of PR-fampridine–treated participantswith clinically meaningful improvements (≥ 15%) [6] in TUG speed (ft/s) vs. placebo ($p = 0.03$). We do not believe the lack of clear effect size for TUG time nullifies the higher proportion of PR-fampridine–treated participantswho were above the clinically meaningful threshold for TUG speed. The effect sizes provide complementary information on the magnitude of effect at a population level, which include participantswith a wide range of mobility. As such, effect sizes are generated based on population-level distribution statistics and are impacted by both inter- and intra-patient variability, which makes it difficult to translate what this type of analysis represents to the individual participantwho is being evaluated for treatment benefit.

The observed means of the MSIS-29 PHYS were very near the scale midpoint, where the scale's ability to detect change is maximal (the scales' 'sweet spot'). The pre-treatment means of the MSWS-12 were above the scale midpoint, skewed towards worse walking ability, which was acceptable as PR-fampridine improves

walking ability, driving the on-treatment scores to the left and to the area of the scale where the ability to detect change is best. As such, the MSWS-12 is well targeted to this sample of walking-disabled participants.

Findings for the BBS and ABILHAND scores were particularly important and are described in the main text. Pre-treatment mean scores of both BBS and ABILHAND were notable for being above the scale midpoint and may have been suboptimal instruments for evaluating the impact of PR-fampridine in this population (for further details see Discussion section in the main text).These features highlight the central role of the COA scales in accurately representing the effects of treatments on how individuals feel and function. The scales are the central dependent variables on which inferences are made, which cannot be underestimated and support the emphasis on providing evidence that COAs are well defined and reliable measures of clinically meaningful concepts of interest in a specific context of use. More focus is needed on developing such instruments, which highlights the value of pilot work in examining the suitability of scales, which cannot just be selected off the shelf.

This study reported effect sizes and standardized change scores to complement the mean change scores, which enable head-to-head comparisons of different instruments and have criteria for interpretation that are widely used. The two calculations give notably different values, which is not surprising as they are conceptually different: Cohen's effect size relates the mean change to the sample variability at baseline, whereas the SRM relates the mean change to the sample variability of change. However, both measurements are interpreted using the same Cohen's criteria. This issue has been raised before, with the recommendation to provide both computations as to provide only one can be misleading [7].

This study also provides valid evidence for a $\geq$ 8-point improvement threshold in MSWS-12 score to indicate a clinically meaningful change. A change of 8 points equates to a Cohen's effect size of 0.40 and an SRM of 0.66, essentially a clinically moderate or 'not insignificant' change [4].

References

1. McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann Neurol. 2001 Jul;50(1):121-7.
2. Polman CH, Wolinsky JS, Reingold SC. Multiple sclerosis diagnostic criteria: three years later. Mult Scler. 2005 Feb;11(1):5-12.
3. Little RA, Rubin DB. Estimation of imputation uncertainty. In Statistical analysis with missing data. 2nd ed. New York: Wiley-Interscience; 2002. p. 75-96.
4. Cohen J. A power primer. Psychol Bull. 1992 Jul;112(1):155-9.
5. Mehta L, McNeill M, Hobart J, Wyrwich KW, Poon J-L, Auguste P, et al. Identifying an important change estimate for the Multiple Sclerosis Walking Scale-12 (MSWS-12v1) for interpreting clinical trial results. Mult Scler J Exp Transl Clin. 2015;1:2055217315596993.
6. Hupperts R, Lycke J, Short C, Gasperini C, McNeill M, Medori R, et al. Prolonged-release fampridine and walking and balance in MS: randomised controlled MOBILE trial. Mult Scler. 2016;22(2):212-21.
7. Middel B, van Sonderen E. Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. Int J Integr Care. 2002 Oct-Dec;2:e15.