ARTICLE

Responsiveness and meaningful improvement of mobility measures following MS rehabilitation

Ilse Baert, PT, PhD, Tori Smedal, PT, PhD, Alon Kalron, PT, PhD, Kamila Rasova, PT, PhD, Adnan Heric-Mansrud, PT, MSc, Rainer Ehling, MD, Iratxe Elorriaga Minguez, PT, Una Nedeljkovic, MD, Andrea Tacchino, PhD, Peter Hellinckx, PT, Greet Adriaenssens, PT, Gosia Stachowiak, PT, Klaus Gusowski, PT, Davide Cattaneo, PT, PhD, Sophie Borgers, PT, Jeffrey Hebert, PT, PhD, Ulrik Dalgas, MSc, PhD, and Peter Feys, PT, PhD

Correspondence
Dr. Baert
ilse.baert@uhasselt.be

Neurology® 2018;00:e1-e13. doi:10.1212/WNL.0000000000006532

Abstract

Objective

To determine responsiveness of functional mobility measures, and provide reference values for clinically meaningful improvements, according to disability level, in persons with multiple sclerosis (pwMS) in response to physical rehabilitation.

Methods

Thirteen mobility measures (clinician- and patient-reported) were assessed before and after rehabilitation in 191 pwMS from 17 international centers (European and United States). Combined anchor- and distribution-based methods were used. A global rating of change scale, from patients' and therapists' perspective, served as external criteria when determining the area under the receiver operating characteristic curve (AUC), the minimally important change (MIC), and the smallest real change (SRC). Patients were stratified into 2 subgroups based on disability level (Expanded Disability Status Scale score ≤4 [n = 72], >4 [n = 119]).

Results

The Multiple Sclerosis Walking Scale–12, physical subscale of the Multiple Sclerosis Impact Scale–29 (especially for the mildly disabled pwMS), Rivermead Mobility Index, and 5-repetition sit-to-stand test (especially for the moderately to severely disabled pwMS) were the most sensitive measures in detecting improvements in mobility. Findings were determined once the AUC (95% confidence interval) was above 0.5, MIC was greater than SRC, and results were comparable from the patient and therapist perspective.

Conclusions

Responsiveness, clinically meaningful improvement, and real changes of frequently used mobility measures were calculated, showing great heterogeneity, and were dependent on disability level in pwMS.

From the REVAL Rehabilitation Research Center (I.B., P.F.), BIOMED Biomedical Research Institute, Faculty of Medicine & Life Sciences, Hasselt University, Belgium; Haukeland University Hospital (Norwegian MS Competence Centre and Department of Physiotherapy) (T.S.), Bergen, Norway; Department of Physical Therapy (A.K.), Sackler Faculty of Medicine, Tel-Aviv University; Multiple Sclerosis Center (A.K.), Sheba Medical Center, Tel-Hashomer, Israel; Department of Rehabilitation, Third Faculty of Medicine (K.R.), Charles University and Faculty Hospital Royal Vineyards, Prague, Czech Republic; Multiple Sclerosis Center (A.H.-M.), Hakadal AS, Norway; Department of Neurology (R.E.), Clinic for Rehabilitation Münster, Austria; Eugenia Epalza Rehabilitation Center (I.E.M.), Bilbao, Spain; Clinical Center in Belgrado (U.N.), Serbia; Italian Multiple Sclerosis Foundation (FISM) (A.T.), Scientific Research Area, Genoa, Italy; National MS Center (P.H.), Melsbroek; AZ Klina (G.A.), Campus De Mick, Rehabilitation, Brasschaat, Belgium; John Paul II Rehabilitation Centre for People with Multiple Sclerosis (G.S.), Borne Sulinowo, Poland; Neurological Rehabilitation Centre Quellenhof (K.G.), Sana AG, Germany; IRCSS Fondazione Don Carlo Gnocchi (D.C.), Milan, Italy; Rehabilitation and MS Center Overpelt (S.B.), Belgium; University of Colorado–Anschutz Medical Campus (J.H.); and Section of Sport Science (U.D.), Department Public Health, Aarhus University, Denmark.

Go to Neurology.org/N for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

Glossary

5STS = 5-Repetition Sit-to-Stand Test; **ABC** = Activities-specific Balance Confidence Scale; **AUC** = area under the curve; **BBS** = Berg Balance Scale; **CI** = confidence interval; **DGI** = Dynamic Gait Index; **EDSS** = Expanded Disability Status Scale; **FSMC** = Fatigue Scale for Motor and Cognitive functions; **FSST** = 4-Square Step Test; **GRS** = global rating of change scale; **MDC_{95}** = minimal detectable change; **MIC** = minimally important change; **MS** = multiple sclerosis; **$MSIS\text{-}29_{phys}$** = physical subscale of the Multiple Sclerosis Impact Scale–29; **MSWS-12** = Multiple Sclerosis Walking Scale–12; **PRO** = patient reported outcome; **pwMS** = people with multiple sclerosis; **RE** = relative efficiency; **RMI** = Rivermead Mobility Index; **ROC** = receiver operating characteristic; **SDMT** = Symbol Digit Modalities Test; **SEM** = standard error of measurement; **SRC** = smallest real change; **TIS-modNV** = Trunk Impairment Scale–modified Norwegian version; **TUG** = Timed Up and Go.

Gait and balance disturbances are common in multiple sclerosis (MS),[1,2] limiting daily life activities[3,4] and increasing fall risk.[5] Therefore, it is important to recognize, quantify, and treat impaired mobility in people with MS (pwMS) using psychometrically sound measures.

Validity and reliability of mobility measures have generally demonstrated good psychometric properties in pwMS.[6–9] However, responsiveness of mobility measures is sparsely researched in MS, focusing on disease progression[10–14] or pharmacologic treatment.[15,16] Only one study[17] investigated responsiveness in response to MS rehabilitation. Importantly, the size of the responsiveness index is affected by the intervention type, methodologic approach, patient characteristics, and timing of data collection.[18] Therefore, head-to-head comparisons of responsiveness of diverse mobility measures in subpopulations of pwMS are important. Specifically, values defining a clinically meaningful change and a real change not attributable to "noise" or "error" are prerequisites for clinical decision-making and research. Such values are reported for only a few mobility measures[10,12,14,16] and rarely after rehabilitation.[17] Consequently, for the majority of mobility measures, values of clinically meaningful improvement after rehabilitation treatment are lacking.

Previously, we investigated[19] responsiveness and clinically meaningful improvement of common walking measures (Timed 25-Foot Walk, 2- and 6-Minute Walk Tests, and Multiple Sclerosis Walking Scale–12) after MS rehabilitation. The present study provides data on mobility measures beyond walking speed/endurance. Purposes of the present study were to (1) compare responsiveness of 13 common functional mobility measures and (2) provide values of real change and clinically meaningful improvement from both patients' and therapists' perspective, according to disability level.

Methods

Participants

In a prospective trial, a convenience sample of 191 pwMS was recruited at 17 centers within the European Rehabilitation in Multiple Sclerosis network for best practice and research in MS rehabilitation (eurims.org), Israel, and the United States: Department of Neurology, Clinic for Rehabilitation Münster, Austria (n = 20); National Multiple Sclerosis Center Melsbroek (n = 14); AZ Klina, Campus De Mick, Rehabilitation, Brasschaat (n = 8); Rehabilitation and Multiple Sclerosis Center Overpelt, Belgium (n = 3); Charles University in Prague and General Faculty Hospital, Czech Republic (n = 26); Neurologic Rehabilitation Centre Quellenhof, Germany (n = 9); Eugenia Epalza Rehabilitation Center Bilbao, Spain (n = 16); Multiple Sclerosis Centre, Sheba Medical Center, Tel-Hashomer, Israel (n = 12); FISM Scientific Research, AISM (n = 11); IRCSS Fondazione Don Carlo Gnocchi Milan, Italy (n = 8); Haukeland University Hospital (Norwegian MS Competence Centre and Department of Physiotherapy) (n = 10); Multiple Sclerosis Center Hakadal AS, Norway (n = 23); John Paul II Rehabilitation Centre for People with Multiple Sclerosis in Borne Sulinowo, Poland (n = 11); Clinical Center in Belgrado, Serbia (n = 17); and University of Colorado—Anschutz Medical Campus (n = 3).

Included participants had a definite diagnosis of MS, Expanded Disability Status Scale score (EDSS) ≥2 and ≤6.5 as determined by neurologists or trained clinicians, no relapse or changes in disease-modifying treatment or no corticoid therapy within the last month, and received at least 10 sessions of physical therapy (inpatient or outpatient rehabilitation), with a maximum duration of 3 months. Exclusion criteria were other medical conditions interfering with mobility (e.g., pregnancy, fractures), other neurologic conditions causing permanent damage (e.g., stroke, Parkinson disease), MS-like syndromes such as neuromyelitis optica, or inability to understand and execute simple instructions.

Standard protocol approvals, registrations, and patient consents

The study was approved by the Ethical Committee of Leuven University Hospital and Hasselt University (Belgium) as well as local ethical committees from each participating center. All participants provided written informed consent. This study's clinical trial registration number is NCT02339688.

Design and outcome measures

A noncontrolled interventional multicenter study design was applied in rehabilitation settings. Age, sex, EDSS, type of MS, and time since diagnosis were recorded at baseline. Cognitive function was assessed by the Symbol Digit Modalities Test (SDMT)[20] and severity of fatigue by the Fatigue Scale for

Motor and Cognitive functions (FSMC).[21] Extent and content of physical rehabilitation (setting, volume, goal, trained activities, and used interventions) were recorded, and will be published separately.

Functional mobility measures were assessed before and after physical rehabilitation, according to a standardized test protocol, including detailed test procedures, verbal instructions, and level of encouragement during testing, besides standardized electronic files to collect and transfer data. Mobility measures were selected based on an unpublished systematic review of psychometric properties of functional mobility measures, recommendations from the Multiple Sclerosis Outcome Measures Taskforce,[22] and discussions within the workgroup. Clinical tests were Timed Up and Go (TUG), $TUG_{cognitive}$, TUG_{manual}, 4-Square Step Test (FSST), 5-Repetition Sit-to-Stand Test (5STS), Berg Balance Scale (BBS), Dynamic Gait Index (DGI), and Trunk Impairment Scale–modified Norwegian version (TIS-modNV). Patient-reported outcome (PRO) measures were Multiple Sclerosis Walking Scale–12 (MSWS-12), Rivermead Mobility Index (RMI), the physical subscale of the Multiple Sclerosis Impact Scale–29 ($MSIS\text{-}29_{phys}$), and the Activities-specific Balance Confidence Scale (ABC). TUG_{manual}, BBS, DGI, TIS-modNV, and ABC were optional and if EDSS was 6.5, DGI and FSST were not assessed given too severe disability to ensure relevant and safe execution.

During TUG, the participant stands up from a chair, walks 3 meters, turns back, and sits down again as quickly and safely as possible while being timed.[23] $TUG_{cognitive}$ involves additionally a cognitive task (subtracting 3 starting from 50). TUG_{manual} involves additionally holding a full cup of water. The FSST assesses dynamic standing balance[24] and involves stepping rapidly forwards, backwards, and sideways to both sides, over a low obstacle (a cane), while being timed. If stepping over a cane was not possible, tape on the floor was used instead consistently. 5STS is a timed test of 5 repetitions of standing up and sitting down as quickly as possible when rising from a chair.[25] In the modified version, persons were allowed to use their arms to push off from a chair with armrests (mod 5STS). BBS is a 14-task scale requiring participants to maintain their balance in positions and tasks of increasing difficulty, such as standing unsupported with eyes closed and picking up an object from the floor from a standing position.[26] DGI assesses a participant's ability to respond to changing task demands during walking, e.g., walking while changing speed and walking over and around obstacles.[27] TIS-modNV is a 6-item ordinal rating scale that evaluates dynamic sitting balance and coordination of trunk movement.[28] MSWS-12 is a PRO of the impact of MS on limitations in one's individual's walking ability.[29] RMI captures 15 activities from turning over in bed to running.[30] MSIS-29 is a 29-item PRO on the impact of MS on day-to-day life in the last 2 weeks, containing 20 items in a physical scale and 9 in a psychological scale.[1] ABC is a 16-item PRO asking people to rate balance confidence in performing everyday activities on a numeric rating scale.[31]

A 7-point Likert-type global rating of change scale (GRS)[32] from both the patient and therapist perspective was applied as external criteria. The following GRS question was asked after the rehabilitation program: "Compared to before rehabilitation, how would you rate your/the participant's walking/moving around and balance now?" The response categories were very much worse, much worse, minimally worse, no change, minimally improved, much improved, very much improved.

Statistical analyses

Descriptive statistics were used. Participant characteristics for those included in the final dataset were compared with the dropouts using Mann-Whitney tests, median tests, or χ^2, as appropriate. As the outcome measures did not show normal distribution, Wilcoxon matched-pairs signed-ranks test and Sign test were used to examine the significance of changes in the mobility measures after physical rehabilitation. In line with our previous research,[19] the differential effect of disability level on the results of responsiveness was investigated, by categorizing participants according to disability as either mild (EDSS ≤ 4) or moderate–severe (EDSS > 4), except for the optional measures due to the small numbers. In all analyses, p values ≤0.05 were considered significant. Data were analyzed with SPSS 24 (SPSS Inc., Chicago, IL) and SAS Enterprise 7.1 (SAS Institute, Cary, NC).

Combined anchor- and distribution-based approach

As in our previous publication,[19] a combination of anchor- and distribution-based methods were used to assess responsiveness.[18] Distribution-based methods, using standardized metrics, focus on an outcome measure's ability to capture the amount of change needed to exceed measurement error, aiming to quantify real change. Anchor-based methods focus on the correspondence of change on the outcome measure of interest with change in an external criterion, aiming to quantify relevant change (i.e., the amount of change that is considered clinically important). Patient- or therapist-rated global rating scales are commonly used as external criteria.

Distribution-based approach

To quantify real change, we calculated the smallest real change (SRC). The SRC takes 2 sources of variability into account: (1) the reliability of the outcome measure and (2) the naturally occurring variability in stable patients. Previously, the SRC was reported at both group and individual level.[33]

$$SRC_{ind} = 1.96 * SD\ (\text{standard deviation of score changes in the stable group})$$

$$SRC_{group} = SRC_{ind}/\sqrt{n}\ (n = \text{number of stable patients})$$

The stable group included patients with response categories minimally worse, no change, or minimally improved on GRS from the patient/therapist perspective.

Outliers for change scores (≤ or ≥3 × SD) in the stable group were omitted.

Anchor-based approach

To examine head-to-head comparisons of responsiveness of diverse mobility measures, we calculated the receiver operating characteristic (ROC) curve and the area under the curve (AUC) with its 95% confidence interval (CI). ROC compares the rate at which the threshold correctly identifies participants' change (sensitivity) with the rate at which participants are identified as showing change in the measure but not in the external criterion (1 – specificity). An AUC (95% CI) value of ≤0.5 indicates that the outcome measure is not responsive (figure). The larger the AUC (the more the ROC curve approaches the upper left corner) the more responsive is the measure. AUC can be interpreted as the probability of accurately discriminating between the patients who improve and those who do not. In our analyses, this discrimination was based on GRS response categories minimally worse, no change, and minimally improved defined as the stable group and much/very much improved as the improved group. The categories of much/very much worse were not retained because of the small percentage of patients perceiving themselves/therapists indicating patients in these categories.

To provide reference values for clinically meaningful improvement, we calculated the minimally important change (MIC), defined as the mean change score in patients who showed a relevant improvement according to the external criterion (much, very much improved). This was calculated by applying linear regression analysis adapted from a previous study[10]:

$$Y\ (t) = \alpha + \beta 1 \times Y\ (t0) + \beta 2 \times \text{improved} + \varepsilon$$

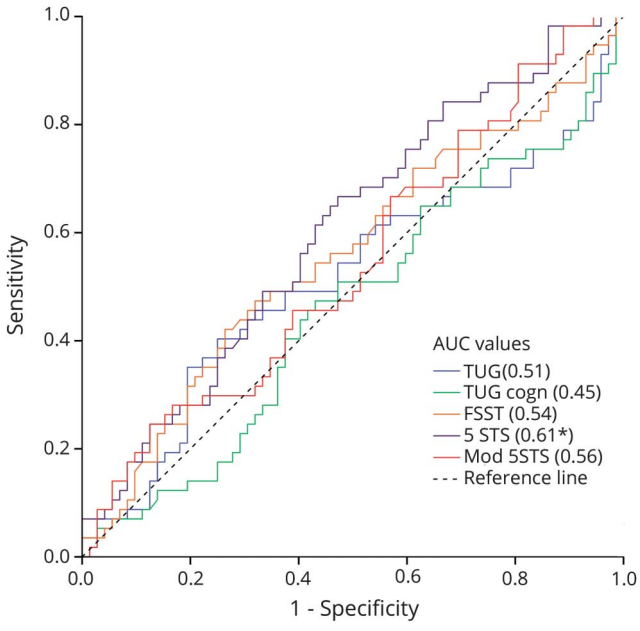
Scores on the mobility measures postintervention were used as dependent variables [Y (t)]. The improved group consisted of participants/therapists perceiving their/the change as very much or much improved on the external criteria, which was transformed as a dummy variable. This dummy variable and the scores on the mobility measures preintervention [Y (t0)] were used as independent variables. The stable group (minimally worse, no change, and minimally improved on the external criteria) was used as the reference group. In the formula, β2 is a differential intercept coefficient, which indicates how much the mean value of the improved group differs from the mean value of the stable group, and provides an estimate for the $MIC_{improvement}$. The distribution of residuals was inspected and outliers removed to get symmetric and homoscedastic distribution.

Weighted kappa (κw) and percentage agreement were calculated to investigate agreement between patient and therapist GRS scores.

Combined anchor- and distribution-based approach

The interaction between SRC and MIC is illustrated and explained in figure 2 in our previous publication.[19] The selection of the most appropriate outcome measure to detect change was based on whether the AUC value (significant) and its 95% CI were above 0.5, whether $|MIC| > |SRC_{ind\ or\ group}|$, and whether the results were comparable from the patient and therapist perspective. Moreover, nonsignificant MIC values means that MIC values of these measures cannot be accurately estimated in this study sample.

Figure Receiver operating characteristic curves (ROCs) and the areas under the ROC (AUC) of clinician-reported functional mobility measures from the patients' perspective for the whole group



5STS = 5-Repetition Sit-to-Stand Test; mod 5STS = modified 5-Repetition Sit-to-Stand Test; FSST = 4-Square Step Test; TUG = Timed Up and Go; TUG_{cogn} = Timed Up and Go cognitive. *p < 0.05 for AUC values.

Data availability

Related study protocols or (supplementary) data will only be shared for secondary analyses on request, including clearly stated data sharing conditions (research question, ownership, security/protection, and confidentiality of data), until 2023.

Results

Sample characteristics

Of the 206 patients who were recruited, 13 were not assessed after rehabilitation due to unforeseen medical circumstances (e.g., relapse, changes in disease-modifying treatment) (n = 5), fewer than 10 therapy sessions (n = 6), or forgotten data collection (n = 2). In addition, for 2 patients, there were missing GRS scores. The 15 excluded patients did not differ significantly from the 191 who remained in the study with respect to age, sex, EDSS, disease duration, type of MS, SDMT, or FSMC. Table 1 shows the participant

characteristics for the whole group and the subgroups based on disability levels. Forty-six percent of participants were inpatient (54% outpatient). On average, they received 28 sessions of therapy (individual, group, or autonomous therapy). Participants demonstrated cognitive dysfunction and severe motor fatigue at group level.

Change after physical rehabilitation

Regardless of disability level, significant improvements on the mobility measures following physical rehabilitation were demonstrated at group level (table 2).

Table 3 details the frequency and percentage of the GRS scores, demonstrating that most of the patients and therapists reported patient's change as minimally improved (37% and 53%, respectively) or much improved (35% and 24%, respectively), while 17% and 15% reported no change, respectively. Agreement between the patient and therapist perspective was moderate (κ_w = 0.46, p < 0.17, 58% agreement).

Sensitivity in detecting improvement after physical rehabilitation

Tables 4 and 5 show the AUCs, MICs, and SRCs for improvement in mobility using the patient/therapist perspective as the external criteria. Table 6 presents a synopsis whether variables met the selection criteria and their interpretation to be sensitive to detect change. 5STS and MSWS-12 met the criteria of significant AUC (95% CI) values > 0.5 (5STS, 0.61 and 0.64; MSWS-12, 0.61 and 0.61, respectively), significant $MIC_{improvement}$ values (5STS, −2.03 seconds and −1.54 seconds; MSWS-12, −8.85 and −6.30, respectively) > $|SRC_{group}|$, and comparable results from the patient and therapist perspective. From a patient perspective, the RMI and $MSIS\text{-}29_{phys}$ also revealed significant AUC (95% CI) values > 0.5 (0.60 and 0.59, respectively) and significant $MIC_{improvement}$ values (0.52 and −4.95, respectively) greater than the real change, as did the BBS (AUC, 0.70 and $MIC_{improvement}$, 2.5) from the therapist perspective. The FSST, RMI, and $MSIS\text{-}29_{phys}$ showed significant $|MIC| > |SRC_{group}|$, but did not show significant AUC values (although borderline for the FSST) from the therapist perspective.

Table 1 Participant characteristics

		Disability subgroups		
Variable	**Whole group**	**Mild (EDSS ≤4)**	**Moderate–severe (EDSS 4.5–6.5)**	***p* Value[a]**
N	191	72	119	
Age, y, mean ± SD	48.7 ± 10.5	47.2 ± 10.3	49.6 ± 10.5	0.12
Sex, n (%)				0.84
Male	68 (36)	25 (35)	43 (36)	
Female	123 (64)	47 (65)	76 (64)	
Living arrangement,[a] n (%)				0.26
Alone	25 (13)	7 (10)	18 (15)	
With family/partner	164 (86)	65 (90)	99 (85)	
EDSS, median (IQR)	5 (3.5–6)	3 (2.5–4)	6 (5–6.5)	<0.01
Type of MS[a]				<0.01
Relapsing-remitting, n (%)	96 (52)	47 (67)	49 (42)	
Secondary progressive, n (%)	51 (27)	7 (10)	44 (38)	
Primary progressive, n (%)	37 (20)	16 (23)	21 (18)	
Years since diagnosis, mean ± SD	11.9 ± 8.1	10.0 ± 6.8	13.1 ± 8.6	0.23
SDMT (0–110), mean ± SD	40 ± 14	43 ± 13	38 ± 15	<0.01
FSMC, mean ± SD				
FSMC total (20–100)	65 ± 18	63 ± 18	66 ± 17	0.55
FSMC cognitve (10–50)	30 ± 11	30 ± 11	30 ± 11	0.97
FSMC motor (10–50)	35 ± 8	34 ± 9	36 ± 8	0.14

Abbreviations: EDSS = Expanded Disability Status Scale; FSMC = Fatigue Scale for Motor and Cognitive Functions; IQR = interquartile range; MS = multiple sclerosis; SDMT = Symbol Digit Modalities Test.
[a] *t* Test, Wilcoxon, median test, or χ^2 between disability subgroups.
[b] Missing cases.

Table 2 Changes in functional mobility outcome measures after physical rehabilitation (mean ± SD or median [IQR])

Outcome meausure	N	Pre	Post	Change (improvement)[a]
Whole group				
TUG, s	185	12.9 ± 7.9	11.9 ± 9.3	−0.9 ± 3.2[b]
TUG cogn, s	180	14.7 ± 9.1	13.3 ± 9.3	−1.1 ± 4.1[b]
FSST, s	164	16.7 ± 11.7	14.0 ± 11.1	−2.2 ± 4.6[b]
5STS, s	162	16.6 ± 7.2	14.6 ± 8.0	−2.3 ± 4.7[b]
mod 5STS, s	156	14.0 ± 5.3	12.9 ± 9.0	−1.8 ± 3.1[b]
TUG manual, s[c]	55	13.7 ± 7.5	12.4 ± 5.4	−1.3 ± 3.8
BBS (0–56)[c]	68	48 (37–52)	49 (39–53)	2 (0–3)[b]
DGI (0–24)[c]	69	15 (12–19)	17 (15–21)	1 (0–3)[b]
TIS-modNV (0–16)[c]	78	10 (7–12)	11 (9–13)	1.5 (0–3)[b]
MSWS-12 (0–100)	191	55.8 ± 27.4	45.0 ± 26.5	−10.8 ± 21.4[b]
RMI (0–14)	191	13 (12–14)	14 (12–15)	0 (0–1)[b]
MSIS-29 phys (0–100)	190	41.5 ± 22.0	32.2 ± 21.1	−9.4 ± 15.8[b]
ABC (0–100)[c]	69	57.2 ± 21.3	61.3 ± 21.4	4.1 ± 13.2[b]
Mild disability group (EDSS ≤ 4)				
TUG, s	72	8.1 ± 3.0	7.4 ± 3.0	−0.7 ± 1.2[b]
TUG cogn, s	69	9.6 ± 4.1	8.5 ± 3.9	−1.1 ± 2.4[b]
FSST, s	70	10.1 ± 3.6	8.9 ± 3.3	−1.3 ± 1.9[b]
5STS, s	70	13.7 ± 6.9	11.4 ± 5.3	−2.2 ± 4.9[b]
mod 5STS, s	65	11.8 ± 4.2	10.2 ± 3.8	−1.6 ± 2.3[b]
MSWS-12	72	40.2 ± 25.5	32.9 ± 24.7	−7.3 ± 19.6[b]
RMI	72	14 (14–15)	15 (14–15)	0 (0–0)[b]
MSIS-29 phys	71	31.8 ± 19.3	24.4 ± 18.2	−7.5 ± 13.8[b]
Moderate–severe disability group (EDSS 4.5–6.5)				
TUG, s	113	15.8 ± 8.6	14.8 ± 10.7	−1.0 ± 4.0[b]
TUG cogn, s	111	17.8 ± 9.9	16.3 ± 10.4	−1.1 ± 4.9[b]
FSST, s	94	21.5 ± 13.1	17.8 ± 13.1	−2.9 ± 5.8[b]
5STS, s	91	18.7 ± 6.8	16.8 ± 8.7	−2.4 ± 4.4[b]
mod 5STS, s	91	15.5 ± 5.4	14.8 ± 10.8	−1.8 ± 3.5[b]
MSWS-12	119	65.2 ± 24.1	52.3 ± 24.9	−12.9 ± 22.3[b]
RMI	119	12 (11–14)	13 (11–14)	0 (0–1)[b]
MSIS-29 phys	119	47.4 ± 21.5	36.9 ± 21.3	−10.5 ± 16.9[b]

Abbreviations: 5STS = 5-Repetition Sit-to-Stand Test; ABC = Activities-specific Balance Confidence Scale; BBS = Berg Balance Scale; DGI = Dynamic Gait Index; EDSS = Expanded Disability Status Scale; FSST = 4-Square Step Test; IQR = interquartile range; mod 5STS = modified 5-Repetition Sit-to-Stand Test; MSWS-12 = Multiple Sclerosis Walking Scale–12; MSIS-29 phys = Multiple Sclerosis Impact Scale–29 physical; RMI = Rivermead Mobility Index; TIS-modNV = Trunk Impairment Scale–modified Norwegian Version; TUG = Timed Up and Go; TUG cogn = Timed Up and Go cognitive; TUG manual = Timed Up and Go manual.
[a] Mobility improvement is indicated by negative change scores on the TUG, TUG cogn, FSST, 5STS, mod 5STS, TUG manual, MSIS-29 phys, but positive change scores on the RMI, BBS, DGI, TIS-modNV, ABC.
[b] $p < 0.05$ by Wilcoxon matched-pairs signed-ranks test or sign test.
[c] Optional measure.

Table 3 Distribution of global rating of change scale from patient's and therapist's perspective on functional mobility

	Perspective from patient, n (%)	Perspective from therapist, n (%)
Very much worse	0 (0)	0 (0)
Much worse	3 (1.6)	2 (1.1)
Minimally worse	4 (2.1)	3 (1.6)
No change	33 (17.3)	29 (15.2)
Minimally improved	70 (36.7)	102 (53.4)
Much improved	66 (34.6)	45 (23.6)
Very much improved	8 (4.2)	8 (4.2)

In the mild disability subgroup, MSWS-12 and MSIS-29_{phys} were the most sensitive in detecting mobility improvement. Significant AUC (95% CI) values (>0.5) and significant $|MIC| > |SRC_{group}|$ were found for MSIS-29_{phys} from the patient perspective (AUC, 0.64 and $MIC_{improvement}$, −6.58) and for MSWS-12 from the therapist perspective (AUC, 0.71 and $MIC_{improvement}$, −12.17). Significant $|MIC| > |SRC_{group}|$, but AUC (95% CI) values < 0.5 were found for MSWS-12 and MSIS-29_{phys}, FSST from the patient and therapist perspective, respectively.

From the patient perspective in the moderate to severe disability subgroup, MSWS-12 showed significant AUC (95% CI) values > 0.5 (0.70) and significant $|MIC| > |SRC_{group}|$ (−7.66), for 5STS and RMI only significant $|MIC| > |SRC_{group}|$ were observed. From the therapist perspective, significant $|MIC| > |SRC_{group}|$ was found for the RMI.

Overall, TUG, TUG_{cogn}, TUG_{manual}, modified 5STS, TIS-modNV, and ABC did not show significant AUC or significant $MIC_{improvement}$ values from either perspective. In addition, none of the outcome measures have a $|MIC| > |SRC_{ind}|$, making them statistically unsuitable for detecting important changes in individuals. However, $MIC_{improvement}$ was mostly greater than SRC_{group}, making them suitable to compare treatment effects at a group level. Mobility measures were mostly more responsive (higher AUC) in mildly disabled persons compared to moderate to severely disabled pwMS.

Discussion

This study aimed to determine which functional mobility measures are sensitive to detect clinically significant changes after physical rehabilitation in a large sample of pwMS. Investigation of head-to-head comparison of responsiveness' magnitude, and provision of real change and clinically meaningful improvement from both the patient and therapist perspective, also considering different disability levels, were performed.

Simultaneous assessment of several mobility measures allows head-to-head comparisons of their responsiveness, enabling evidence-based choices regarding selection of mobility measures. Our results suggest that the patient-reported MSWS-12, the MSIS-29_{phys} (especially for the mildly disabled pwMS), RMI (especially for the moderate to severely disabled pwMS), and capacity test 5STS (especially for the moderate to severely disabled pwMS) were the most sensitive measures in detecting improvements in functional mobility after physical rehabilitation. Provision of reference values for clinically meaningful improvement and real change after physical rehabilitation, established in our studies, can guide clinicians and researchers in interpreting if a change in mobility ability exceeds measurement variability (SRC values) and whether this change is meaningful (MIC values). Other psychometric properties (reliability, validity, internal consistency, floor/ceiling effect, normative data, and clinical utility) have been extensively evaluated for the MSWS-12 and MSIS-29_{phys}, with demonstration of adequate to excellent results.[6,7,13,22] However, reliability and validity of the RMI[34,35] and 5STS[36] was less often investigated in the MS population.

On the other hand, although TUG,[35,37–41] TUG_{cogn},[41,42] and ABC[35,37,38,42,43] are appropriate, reliable, and valid in the MS population, our results revealed they are not sufficiently sensitive in detecting clinically relevant changes after rehabilitation. This may be due to the small delta values for these outcome measures that were found in the present study.

Studies providing reference values for clinically meaningful change in diverse functional mobility measures associated with improved function after physical rehabilitation in pwMS are rare. Previously, one study[17] reported a minimal clinically important difference of +3 points and an AUC value of 0.65 for the BBS, in response to rehabilitation. The present study revealed similar results (MIC 2.5 and AUC 0.70) from the therapist perspective, albeit not from the patient perspective. This is likely explained by larger delta values of the BBS in a previous study.[17] Another study[14] reported AUC and MIC for deterioration of a modified BBS. In the present study, smaller values were found from the patient perspective, while larger values from the therapist perspective were observed. For RMI, a previous study[10] reported AUC values of 0.67 and 0.65 from the patient and therapist perspective in mildly disabled pwMS, which is somewhat higher than our results. Our absolute values of MIC and SRC were also smaller. However, it can be questioned whether it is of relevance to directly compare MIC values of improvement after short-lasting rehabilitation interventions with values of deterioration derived with a perspective of years. For MSIS-29_{phys}, AUC values of 0.68 and 0.60 from the patient and physician perspective after steroid treatment were reported,[15] which are slightly higher than our results. The different anchors, timeframes, types of interventions vs monitoring deterioration, and sample characteristics applied in these studies may have influenced the responsiveness indexes.

Table 4 Combined anchor- and distribution-based responsiveness (AUC, MIC, SRC) for improvement on functional mobility using patient's perspective as external criteria

	ROC			$MIC_{improved}$			SRC	
Outcome meausure	**AUC**	**95% CI**	***p* Value**	**β_2**	**SE**	***p* Value**	**SRC_{ind}**	**SRC_{group}**
Whole group								
TUG, s	0.51	0.43–0.60	0.76	−0.02	0.49	0.97	−4.47	−0.44
TUG cogn, s	0.45	0.36–0.54	0.28	0.96	0.65	0.14	−5.19	−0.52
FSST, s	0.54	0.45–0.63	0.41	−1.02	0.71	0.15	−6.31	−0.67
5STS, s	**0.61**	0.52–0.70	0.02	**−2.03**	0.68	<0.01	−7.35	−0.75
mod 5STS, s	0.56	0.47–0.65	0.20	−0.71	0.40	0.08	−5.65	−0.60
TUG manual, s	0.47	0.32–0.63	0.72	−0.82	0.72	0.26	−5.37	−1.10
BBS	0.53	0.39–0.67	0.71	**0.87**	0.04	<0.01	6.18	1.15
DGI	0.36	0.22–0.49	0.04	**−1.36**	0.63	0.04	4.25	0.80
TIS-modNV	0.49	0.35–0.62	0.83	−0.28	0.61	0.65	5.29	0.87
MSWS-12	**0.61**	0.52–0.70	0.02	**−8.85**	2.79	<0.01	−31.31	−3.04
RMI	**0.60**	0.51–0.68	0.03	**0.52**	0.18	<0.01	2.55	0.25
MSIS-29 phys	**0.59**	0.50–0.67	0.05	**−4.95**	2.13	0.02	−22.72	−2.22
ABC	0.56	0.42–0.70	0.39	4.00	2.60	0.13	22.69	4.14
Mild disability group (EDSS ≤ 4)								
TUG, s	0.57	0.43–0.71	0.32	−0.27	0.29	0.35	−2.10	−0.34
TUG cogn, s	0.46	0.31–0.60	0.56	0.46	0.56	0.41	−3.69	−0.60
FSST, s	0.59	0.45–0.73	0.21	−0.43	0.44	0.33	−2.97	−0.48
5STS, s	0.62	0.51–0.74	0.05	−1.44	0.98	0.15	−5.53	−0.89
mod 5STS, s	0.54	0.40–0.69	0.56	−0.57	0.53	0.28	−3.76	−0.61
MSWS-12	0.56	0.45–0.68	0.27	**−11.11**	4.25	0.01	−27.48	−4.46
RMI	0.63	0.49–0.77	0.07	0.23	0.16	0.15	0.90	0.14
MSIS-29 phys	**0.64**	0.50–0.79	0.05	**−6.58**	3.07	0.04	−14.06	−2.25
Moderate-severe disability group (EDSS 4.5–6.5)								
TUG, s	0.48	0.37–0.59	0.75	0.18	0.78	0.81	−5.42	−0.67
TUG cogn, s	0.45	0.34–0.56	0.38	1.36	0.98	0.17	−5.94	−0.75
FSST, s	0.52	0.40–0.64	0.79	−1.35	1.21	0.27	−7.68	−1.09
5STS, s	0.58	0.44–0.72	0.25	**−2.18**	0.92	0.02	−8.42	−1.13
mod 5STS, s	0.56	0.45–0.68	0.30	−0.84	0.57	0.14	−6.74	−0.94
MSWS-12	**0.70**	0.58–0.83	0.05	**−7.66**	3.71	0.04	−32.84	−3.98
RMI	0.58	0.48–0.69	0.13	**0.69**	0.27	0.02	3.14	0.38
MSIS-29 phys	0.56	0.45–0.67	0.31	−3.85	2.91	0.19	−26.46	−3.26

Abbreviations: 5STS = 5-Repetition Sit-to-Stand Test; ABC = Activities-specific Balance Confidence Scale; AUC = area under the receiver operating characteristic curve; β2 = differential intercept coefficient; BBS = Berg Balance Scale; CI = confidence interval; DGI = Dynamic Gait Index; EDSS = Expanded Disability Status Scale; FSST = 4-Square Step Test; ind = individual; MIC = minimally important change; mod TUG manual = Timed Up and Go manual; MSIS-29 phys = Multiple Sclerosis Impact Scale–29 physical; MSWS-12 = Multiple Sclerosis Walking Scale–12; RMI = Rivermead Mobility Index; ROC = receiver operating characteristic; SE = standard error; SRC = smallest real change; TUG = Timed Up and Go; TUG cogn = Timed Up and Go cognitive; TIS-modNV = Trunk Impairment Scale–modified Norwegian version.
Values are significant for AUC (95% CI) values if >0.5.

Table 5 Combined anchor- and distribution-based responsiveness (AUC, MIC, SRC) for improvement on functional mobility using therapist's perspective as external criteria

	ROC			$MIC_{improved}$			SRC	
Outcome meausure	**AUC**	**95% CI**	***p* Value**	**β_2**	**SE**	***p* Value**	**SRC_{ind}**	**SRC_{group}**
Whole group								
TUG, s	0.59	0.49–0.69	0.06	−0.75	0.53	0.16	−4.82	−0.42
TUG cogn, s	0.56	0.46–0.66	0.25	−0.64	0.71	0.37	−6.30	−0.56
FSST, s	0.63	0.52–0.73	0.05	**0.90**	0.03	<0.01	−7.30	−0.69
5STS, s	**0.64**	0.54–0.74	<0.01	**−1.54**	0.77	0.05	−9.70	−0.89
mod 5STS, s	0.55	0.45–0.66	0.32	−0.34	0.44	0.43	−4.53	−0.44
TUG manual, s	0.59	0.43–0.76	0.25	−0.28	0.76	0.71	−3.74	−0.65
BBS	**0.70**	0.55–0.84	<0.01	**2.50**	1.12	0.03	6.13	0.97
DGI	0.60	0.45–0.74	0.18	0.59	0.66	0.38	3.99	0.63
TIS-modNV	0.52	0.39–0.66	0.73	−0.21	0.64	0.74	5.94	0.86
MSWS-12	**0.61**	0.51–0.70	0.03	**−6.30**	3.14	0.05	−34.81	−3.03
RMI	0.53	0.43–0.63	0.49	**0.69**	0.04	<0.01	2.48	0.22
MSIS-29 phys	0.58	0.48–0.67	0.10	**−4.84**	2.31	0.04	−26.18	−2.28
ABC	0.55	0.42–0.69	0.45	1.96	2.68	0.47	22.83	3.57
Mild disability group (EDSS≤4)								
TUG, s	0.64	0.45–0.84	0.10	−0.63	0.34	0.07	−2.16	−0.29
TUG cogn, s	0.52	0.33–0.72	0.79	−0.22	0.70	0.75	−3.48	−0.47
FSST, s	0.62	0.45–0.79	0.17	**0.83**	0.06	<0.01	−3.31	−0.45
5STS, s	0.66	0.49–0.82	0.09	−1.80	1.19	0.13	−10.22	−1.38
mod 5STS, s	0.57	0.38–0.75	0.48	−1.10	0.65	0.10	−4.17	−0.58
MSWS-12	**0.71**	0.56–0.86	0.02	**−12.17**	5.22	0.03	−30.37	−4.10
RMI	0.56	0.39–0.74	0.46	**0.67**	0.05	<0.01	1.40	0.19
MSIS-29 phys	0.63	0.44–0.82	0.13	**−11.94**	3.48	<0.01	−18.49	−2.49
Moderate–severe disability group (EDSS 4.5–6.5)								
TUG, s	0.57	0.45–0.69	0.25	−0.84	0.81	0.31	−6.14	−0.72
TUG cogn, s	0.57	0.45–0.69	0.25	−0.96	1.03	0.36	−7.75	−0.91
FSST, s	0.61	0.49–0.74	0.09	**0.91**	0.05	<0.01	−9.40	−1.22
5STS, s	0.62	0.49–0.75	0.08	−1.75	0.98	0.08	−9.30	−1.17
mod 5STS, s	0.54	0.41–0.66	0.56	−0.04	0.6	0.94	−4.87	−0.66
MSWS-12	0.55	0.44–0.67	0.34	−4.66	3.97	0.24	−37.48	−4.27
RMI	0.52	0.40–0.64	0.74	**0.63**	0.06	<0.01	3.03	0.35
MSIS-29 phys	0.54	0.42–0.65	0.54	−2.31	3.02	0.44	−30.28	−3.45

Abbreviations: 5STS = 5-Repetition Sit-to-Stand Test; ABC = Activities-specific Balance Confidence Scale; AUC = area under the receiver operating characteristic curve; β2 = differential intercept coefficient; BBS = Berg Balance Scale; CI = confidence interval; DGI = Dynamic Gait Index; EDSS = Expanded Disability Status Scale; FSST = 4-Square Step Test; ind = individual; MIC = minimally important change; mod 5STS = modified 5-Repetition Sit-to-Stand Test; mod TUG manual = Timed Up and Go manual; MSIS-29 phys = Multiple Sclerosis Impact Scale–29 physical; MSWS-12 = Multiple Sclerosis Walking Scale–12; RMI = Rivermead Mobility Index; ROC = receiver operating characteristic; SE = standard error; SRC = smallest real change; TUG = Timed Up and Go; TUG cogn = Timed Up and Go cognitive; TIS-modNV = Trunk Impairment Scale–modified Norwegian version
Values are significant for AUC (95% CI) values if >0.5.

Table 6 Synopsis whether variables did or did not meet the selection criteria and their interpretation to be sensitive to detect change

Outcome measure	Significant AUC (95% CI) > 0.5		Significant $MIC_{improved}$		$\|MIC\| > \|SRC_{group}\|$[a]		Similar results patient and therapist perspective	Interpretation sensitivity (at group level)
	Patient perspective	Therapist perspective	Patient perspective	Therapist perspective	Patient perspective	Therapist perspective		
Whole group								
TUG, s						x	Mostly	Insensitive
TUG cogn, s					x	x	Yes	Insensitive
FSST, s				x	x	x	Mostly	Insensitive
5STS, s	x	x	x	x	x	x	Yes	Sensitive
mod 5STS, s					x		Mostly	Insensitive
TUG manual, s[b]							Yes	Insensitive
BBS[b]		x	x	x		x	No	Indeterminate
DGI[b]			x		x		No	Indeterminate
TIS-modNV[b]							Yes	Insensitive
MSWS-12	x	x	x	x	x	x	Yes	Sensitive
RMI	x		x	x	x	x	Mostly	Sensitive
MSIS-29 phys	x		x	x	x	x	Mostly	Sensitive
ABC[b]							Yes	Insensitive
Mild disability group (EDSS ≤ 4)								
TUG, s						x	Mostly	Insensitive
TUG cogn, s							Yes	Insensitive
FSST, s				x		x	No	Insensitive
5STS, s					x	x	Yes	Insensitive
mod 5STS, s						x	Mostly	Insensitive
MSWS-12		x	x	x	x	x	Mostly	Sensitive
RMI				x	x	x	Mostly	Insensitive
MSIS-29 phys	x		x	x	x	x	Mostly	Sensitive
Moderate-severe disability group (EDSS 4.5–6.5)								
TUG, s						x	Mostly	Insensitive
TUG cogn, s					x	x	Yes	Insensitive
FSST, s				x	x		No	Insensitive
5STS, s		x	x		x	x	No	Sensitive, to be confirmed
mod 5STS, s							Yes	Insensitive
MSWS-12	x		x		x	x	No	Sensitive, to be confirmed

Continued

Table 6 Synopsis whether variables did or did not meet the selection criteria and their interpretation to be sensitive to detect change *(continued)*

Outcome meausure	Significant AUC (95% CI) > 0.5		Significant $MIC_{improved}$		$\|MIC\| > \|SRC_{group}\|$[a]		Similar results patient and therapist perspective	Interpretation sensitivity (at group level)
	Patient perspective	Therapist perspective	Patient perspective	Therapist perspective	Patient perspective	Therapist perspective		
RMI			x	x	x	x	Yes	Indeterminate
MSIS-29 phys					x		Mostly	Insensitive

Abbreviations: 5STS = 5-Repetition Sit-to-Stand Test; ABC = Activities-specific Balance Confidence Scale; AUC = area under the receiver operating characteristic curve; BBS = Berg Balance Scale; CI = confidence interval; DGI = Dynamic Gait Index; EDSS = Expanded Disability Status Scale; FSST = 4-Square Step Test; MIC = minimally important change; mod 5STS = modified 5-Repetition Sit-to-Stand Test; MSIS-29 phys = Multiple Sclerosis Impact Scale–29 physical; MSWS-12 = Multiple Sclerosis Walking Scale–12; RMI = Rivermead Mobility Index; SRC = smallest real change; TIS-modNV = Trunk Impairment Scale–modified Norwegian version; TUG = Timed Up and Go; TUG cogn = Timed Up and Go cognitive; TUG manual = Timed Up and Go manual.
[a] None of the outcome measures has a |MIC| > |SRCind|, making them statistically unsuitable for detecting important changes in individuals.
[b] Optional measure.

In symptomatic treatment studies with fampridine,[16,44,45] a clinically meaningful improvement on the MSWS-12 was proposed to be 4-, 6-, and 8-point reduction, which corresponds to the values found in the present study, being −8.9 and −6.3 from the patient and therapist perspective, respectively. Our previous publication[19] revealed values of $MIC_{improvement}$ after rehabilitation of −10.4 and −11.4 from the patient and therapist perspective, respectively. Due to diversity in sample, treatment, anchor, and calculation method, no clear consensus on a standard clinically meaningful improvement of the MSWS-12 can be made so far, but taken together it seems that the value ranges between −6 and −11 points.

Our results revealed that variability in mobility measures increases with disability and therefore might explain that mobility measures were mostly more responsive in mildly disabled persons compared to moderate to severely disabled pwMS.

We found an agreement between patient and therapist perspective of 58%. Participants could shift between the stable and improved group depending on the perspective (patient vs therapist) and affect the values of AUC, MIC, and SRC. However, no consistency across differences could be found. Also a recent study[46] revealed differences between patients and physicians in value statements of bodily functions in MS. Although both perspectives are of clinical importance, similarity in responsiveness indexes (AUC [95% IC] > 0.5 and significant $|MIC| > |SRC_{group}|$) increases our confidence in conclusions of sensitive outcome measures.

So far, no consensus exists on how to report responsiveness. In the MS literature,[29,40,47,48] (relative) responsiveness is often based on a distribution based approach, reporting values as standardized response mean, standard error of measurement (SEM), relative efficiency (RE), and minimal detectable change (MDC_{95}). However, these indexes only take the measurement error of the outcome measure into account, while ignoring the naturally occurring variability in stable patients. In contrast, the SRC considers both aspects, suggesting this to be a superior expression of responsiveness. Comparing present SRC values with previously reported SRC or MDC_{95} (SEM) values in longitudinal[10–14] or reliability studies[40,47,48] revealed that our SRC_{group} values were smaller. For example, one study[12] reported a MDC_{95} of 22 for the MSWS-12, while another study[11] reported 6.6. The latter is more in line with our current and previous[19] SRC_{group} values (−3 to −4.6). Further, a MDC_{95} of −7.2 was found following fampridine treatment.[16]

Some methodologic considerations are warranted. To calculate the MIC values, the improved group was used, instead of the minimally improved group. In our initial analyses, the category minimally improved was used for calculation of $MIC_{improvement}$ as this seems, at least semantically, the most appropriate to determine the MIC. However, hardly significant values for MIC and SRC could be found and were therefore not useful. Consequently, the group improved was defined based on response categories much improved and very much improved of the GRS and also the stable group was adapted. This is in line with publications[49,50] on responsiveness in other populations. There is a possibility that the gap between the response category minimally improved and much improved is too large. Therefore, adding an extra response category of improved to the GRS scale may be worthwhile.

A detailed inspection of raw data revealed that the scoring on the EDSS did not always intuitively correspond with scoring on the mobility performance scales or used walking aids (e.g., 7 of the 72 [9.7%] persons with an EDSS ≤4 report using a cane/crutch). An explanation might be that pwMS use walking aids for outdoor mobility, without strictly needing it. Therefore, subgroups stratification based on EDSS score is justified. Some $MIC_{improvement}$ values indicated deterioration rather than improvement (e.g., TUG_{cogn} from the patient perspective and FSST from the therapist perspective). No

plausible explanation could be found for this finding. Furthermore, some of the measures were optional, resulting in a smaller sample size that may have influenced a few of the head-to-head comparisons. In fact, sensitivity of the optional measures may now give uncertain estimates due to the low number. Further research including a larger sample size is required to enhance the precision of estimates that might unmask significance due to stronger analyses power.

This study has several strengths. A wide range of mobility measures were assessed simultaneously, allowing a head-to-head comparison of their responsiveness. Physical rehabilitation was expected to cause improvement in applied outcomes across the disability spectrum, which was demonstrated in the study, supporting the applied approach of examining responsiveness of mobility measures within this context. This study allowed inpatient and outpatient rehabilitation programs, meaning that the results are relevant to both settings. The large sample allowed for documenting values of real change and clinically meaningful improvement with differentiation in subgroups with mild vs moderate disability. $MIC_{improvement}$ was estimated from the differential intercept coefficient, which is the difference between the mean value of the improved group compared to the mean value of the stable group, and therefore not simply based on change scores across GRS categories. To calculate real change, not only reliability of the outcome measure but also the naturally occurring variability in stable patients were taken into account. Our selection of responsive outcome measures was based on 3 criteria.

The present study results indicate that the patient-reported MSWS-12, the $MSIS\text{-}29_{phys}$ (especially for mildly disabled pwMS), RMI (especially for moderate to severely disabled pwMS), and capacity scale 5STS (especially for moderate to severely disabled pwMS) are the most sensitive measures in detecting improvements in mobility after physical rehabilitation. Values of real change and clinically meaningful improvement are provided for interpreting research findings and when evaluating mobility performance in clinical practice. Still, further research should confirm our reference values.

Author contributions

All authors and contributors substantially contributed, including conception or design of the study, analysis or interpretation of the data, or drafting or revising the manuscript for intellectual content. I. Baert: conception or design of the study, analysis or interpretation of the data, drafting or revising the manuscript for intellectual content. T. Smedal: conception or design of the study, analysis or interpretation of the data, drafting or revising the manuscript for intellectual content. A. Kalron: conception or design of the study, analysis or interpretation of the data, drafting or revising the manuscript for intellectual content. K. Rasova: revising the manuscript for intellectual content. A. Heric-Mansrud: revising the manuscript for intellectual content. R. Ehling: revising the manuscript for intellectual content. I.E. Minguez: revising the manuscript for intellectual content. U. Nedeljkovic: revising the manuscript for intellectual content. A. Tacchino: revising the manuscript for intellectual content. P. Hellinckx: revising the manuscript for intellectual content. G. Adriaenssens: revising the manuscript for intellectual content. G. Stachowiak: revising the manuscript for intellectual content. K. Gusowski: revising the manuscript for intellectual content. D. Cattaneo: revising the manuscript for intellectual content. S. Borgers: revising the manuscript for intellectual content. J. Hebert: revising the manuscript for intellectual content. U. Dalgas: conception or design of the study, analysis or interpretation of the data, drafting or revising the manuscript for intellectual content. P. Feys: conception or design of the study, analysis or interpretation of the data, drafting or revising the manuscript for intellectual content.

Acknowledgment

The European Rehabilitation in Multiple Sclerosis network (eurims.org) is acknowledged for facilitating inter-European consultation and testing. The authors thank all the participating persons with MS and the staff of each participating site for cooperation and Dr. I. Lamers for discussion of the responsiveness approach.

Study funding

This study was partially funded via restricted educational grant from Novartis Pharma AG to Rehabilitation in Multiple Sclerosis, the European network for best practice and research in MS rehabilitation (eurims.org).

Disclosure

I. Baert and T. Smedal report no disclosures relevant to the manuscript. A. Kalron has received research support, travel grants, and teaching honoraria from Idec and Merck Serono. K. Rasova, A. Heric-Mansrud, R. Ehling, I. Elorriaga Minguez, U. Nedeljkovic, A. Tacchino, P. Hellinckx, G. Adriaenssens, G. Stachowiak, K. Gusowski, D. Cattaneo, S. Borgers, and J. Hebert report no disclosures relevant to the manuscript. U. Dalgas has received research support, travel grants, and teaching honoraria from Biogen Idec, Merck Serono, and Genzyme/Sanofi Aventis, and has served as PI for a Biogen-sponsored multicenter study (ACTIMS). P. Feys is steering committee member of Neurocompass, participated in advisory board meetings of Biogen Idec, and received teaching honoraria for EXCEMED and PARADIGMS. Go to Neurology.org/N for full disclosures.

Publication history

Received by *Neurology* March 5, 2018. Accepted in final form August 2, 2018.

References

1. Hobart J, Lamping D, Fitzpatrick R, Riazi A, Thompson A. The Multiple Sclerosis Impact Scale (MSIS-29): a new patient-based outcome measure. Brain 2001;124:962–973.
2. Soyuer F, Mirza M, Erkorkmaz U. Balance performance in three forms of multiple sclerosis. Neurol Res 2006;28:555–562.
3. Heesen C, Bohm J, Reich C, Kasper J, Goebel M, Gold SM. Patient perception of bodily functions in multiple sclerosis: gait and visual function are the most valuable. Mult Scler 2008;14:988–991.
4. Larocca NG. Impact of walking impairment in multiple sclerosis: perspectives of patients and care partners. Patient 2011;4:189–201.

5. Sosnoff JJ, Socie MJ, Boes MK, et al. Mobility, balance and falls in persons with multiple sclerosis. PLoS One 2011;6:e28021.
6. Bethoux F, Bennett S. Evaluating walking in patients with multiple sclerosis: which assessment tools are useful in clinical practice? Int J MS Care 2011;13:4–14.
7. Kieseier BC, Pozzilli C. Assessing walking disability in multiple sclerosis. Mult Scler 2012;18:914–924.
8. Tyson S, Connell L. The psychometric properties and clinical utility of measures of walking and mobility in neurological conditions: a systematic review. Clin Rehabil 2009;23:1018–1033.
9. Tyson SF, Connell LA. How to measure balance in clinical practice: a systematic review of the psychometrics and clinical utility of measures of balance activity for neurological conditions. Clin Rehabil 2009;23:824–840.
10. de Groot V, Beckerman H, Uitdehaag BMJ, et al. The usefulness of evaluative outcome measures in patients with multiple sclerosis. Brain 2006;129:2648–2659.
11. Freeman J, Walters R, Ingram W, Slade A, Hobart J, Zajicek J. Evaluating change in mobility in people with multiple sclerosis: relative responsiveness of four clinical measures. Mult Scler 2013;19:1632–1639.
12. Learmonth YC, Dlugonski DD, Pilutti LA, Sandroff BM, Motl RW. The reliability, precision and clinically meaningful change of walking assessments in multiple sclerosis. Mult Scler 2013;19:1784–1791.
13. McGuigan C, Hutchinson M. The Multiple Sclerosis Impact Scale (MSIS-29) is a reliable and sensitive measure. J Neurol Neurosurg Psychiatry 2004;75:266–269.
14. Paltamaa J, Sarasoja T, Leskinen E, Wikstrom J, Malkia E. Measuring deterioration in international classification of functioning domains of people with multiple sclerosis who are ambulatory. Phys Ther 2008;88:176–190.
15. Giordano A, Pucci E, Naldi P, et al. Responsiveness of patient reported outcome measures in multiple sclerosis relapses: the REMS study. J Neurol Neurosurg Psychiatry 2009;80:1023–1028.
16. Mehta L, McNeill M, Hobart J, 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.
17. Gervasoni E, Jonsdottir J, Montesano A, Cattaneo D. Minimal clinically important difference of Berg Balance Scale in people with multiple sclerosis. Arch Phys Med Rehabil 2017;98:337–340 e2.
18. Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. Phys Ther 2006;86:735–743.
19. 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. Neurorehabil Neural Repair 2014;28:621–631.
20. Smith A. Symbol Digit Modalities Test: Manual. Los Angeles: Western Psychological Corporation; 1982.
21. Penner IK, Raselli C, Stocklin M, Opwis K, Kappos L, Calabrese P. The Fatigue Scale for Motor and Cognitive Functions (FSMC): validation of a new instrument to assess multiple sclerosis-related fatigue. Mult Scler 2009;15:1509–1517.
22. Potter K, Cohen ET, Allen DD, et al. Outcome measures for individuals with multiple sclerosis: recommendations from the American Physical Therapy Association Neurology Section task force. Phys Ther 2014;94:593–608.
23. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991;39:142–148.
24. Dite W, Temple VA. A clinical test of stepping and change of direction to identify multiple falling older adults. Arch Phys Med Rehabil 2002;83:1566–1571.
25. McCarthy EK, Horvat MA, Holtsberg PA, Wisenbaker JM. Repeated chair stands as a measure of lower limb strength in sexagenarian women. J Gerontol A Biol Sci Med Sci 2004;59:1207–1212.
26. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Can J Public Health 1992;83(suppl 2):S7–S11.
27. Shumway-Cook A, Taylor CS, Matsuda PN, Studer MT, Whetten BK. Expanding the scoring system for the dynamic gait index. Phys Ther 2013;93:1493–1506.
28. Gjelsvik B, Breivik K, Verheyden G, Smedal T, Hofstad H, Strand LI. The Trunk Impairment Scale: modified to ordinal scales in the Norwegian version. Disabil Rehabil 2012;34:1385–1395.
29. Hobart JC, Riazi A, Lamping DL, Fitzpatrick R, Thompson AJ. Measuring the impact of MS on walking ability: the 12-Item MS Walking Scale (MSWS-12). Neurology 2003;60:31–36.
30. Collen FM, Wade DT, Robb GF, Bradshaw CM. The Rivermead Mobility Index: a further development of the Rivermead motor assessment. Int Disabil Stud 1991;13:50–54.
31. Powell LE, Myers AM. The Activities-specific Balance Confidence (ABC) scale. J Gerontol A Biol Sci Med Sci 1995;50A:M28–M34.
32. Guyatt GH, Norman GR, Juniper EF, Griffith LE. A critical look at transition ratings. J Clin Epidemiol 2002;55:900–908.
33. Pfennings LE, van der Ploeg HM, Cohen L, Polman CH. A comparison of responsiveness indices in multiple sclerosis patients. Qual Life Res 1999;8:481–489.
34. Steen Krawcyk R, Hagell P, Sjodahl Hammarlund C. Danish translation and psychometric testing of the Rivermead Mobility Index. Acta Neurol Scand 2013;128:e20–25.
35. Gijbels D, Alders G, Van HE, et al. Predicting habitual walking performance in multiple sclerosis: relevance of capacity and self-report measures. Mult Scler 2010;16:618–626.
36. Moller AB, Bibby BM, Skjerbaek AG, et al. Validity and variability of the 5-repetition sit-to-stand test in patients with multiple sclerosis. Disabil Rehabil 2012;34:2251–2258.
37. Cattaneo D, Regola A, Meotti M. Validity of six balance disorders scales in persons with multiple sclerosis. Disabil Rehabil 2006;28:789–795.
38. Cavanaugh JT, Gappmaier VO, Dibble LE, Gappmaier E. Ambulatory activity in individuals with multiple sclerosis. J Neurol Phys Ther 2011;35:26–33.
39. Kalron A, Dolev M, Givon U. Further construct validity of the timed up-and-go test as a measure of ambulation in multiple sclerosis patients. Eur J Phys Rehabil Med 2017;53:841–847.
40. Learmonth YC, Paul L, McFadyen AK, Mattison P, Miller L. Reliability and clinical significance of mobility and balance assessments in multiple sclerosis. Int J Rehabil Res 2012;35:69–74.
41. Nilsagard Y, Lundholm C, Gunnarsson LG, Dcnison E. Clinical relevance using timed walk tests and "timed up and go" testing in persons with multiple sclerosis. Physiother Res Int 2007;12:105–114.
42. Nilsagard Y, Carling A, Forsberg A. Activities-specific balance confidence in people with multiple sclerosis. Mult Scler Int 2012;2012:613925.
43. Cattaneo D, Jonsdottir J, Repetti S. Reliability of four scales on balance disorders in persons with multiple sclerosis. Disabil Rehabil 2007;29:1920–1925.
44. Hobart J, Blight AR, Goodman A, Lynn F, Putzki N. Timed 25-foot walk: direct evidence that improving 20% or greater is clinically meaningful in MS. Neurology 2013;80:1509–1517.
45. Hobart J. Prolonged-release fampridine for multiple sclerosis: was the effect on walking ability clinically significant. Mult Scler 2010;16:S172.
46. Heesen C, Haase R, Melzig S, et al. Perceptions on the value of bodily functions in multiple sclerosis. Acta Neurol Scand 2018;137:356–362.
47. Paltamaa J, West H, Sarasoja T, Wikstrom J, Malkia E. Reliability of physical functioning measures in ambulatory subjects with MS. Physiother Res Int 2005;10:93–109.
48. Wagner JM, Norris RA, Van Dillen LR, Thomas FP, Naismith RT. Four square step test in ambulant persons with multiple sclerosis: validity, reliability, and responsiveness. Int J Rehabil Res 2013;36:253–259.
49. Maughan EF, Lewis JS. Outcome measures in chronic low back pain. Eur Spine J 2010;19:1484–1494.
50. Golicki D, Niewada M, Karlinska A, et al. Comparing responsiveness of the EQ-5D-5L, EQ-5D-3L and EQ VAS in stroke patients. Qual Life Res 2015;24:1555–1563.