2014•2015

FACULTEIT GENEESKUNDE EN LEVENSWETENSCHAPPEN

master in de revalidatiewetenschappen en de kinesitherapie

Masterproef

Validity of mobility measures in multiple sclerosis, according disability level: a RIMS multi-center study.

Promotor :
Prof. dr. Peter FEYS

Copromotor :
dr. Ilse BAERT

Laura Jans , Kristien Vanoppen

Scriptie ingediend tot het behalen van de graad van master in de revalidatiewetenschappen en de kinesitherapie

universiteit hasselt
KNOWLEDGE IN ACTION

Universiteit Hasselt | Campus Hasselt | Martelarenlaan 42 | BE-3500 Hasselt
Universiteit Hasselt | Campus Diepenbeek | Agoralaan Gebouw D | BE-3590 Diepenbeek

universiteit hasselt
KNOWLEDGE IN ACTION

2014•2015

FACULTEIT GENEESKUNDE EN LEVENSWETENSCHAPPEN

master in de revalidatiewetenschappen en de kinesitherapie

Masterproef

Validity of mobility measures in multiple sclerosis, according disability level: a RIMS multi-center study.

Promotor :
Prof. dr. Peter FEYS

Copromotor :
dr. Ilse BAERT

Laura Jans , Kristien Vanoppen

Scriptie ingediend tot het behalen van de graad van master in de revalidatiewetenschappen en de kinesitherapie

universiteit hasselt

KNOWLEDGE IN ACTION

Validity of mobility measures in multiple sclerosis, according disability level: a RIMS multi-center study.

“According to the guidelines of the 'Multiple Sclerosis Journal', which conforms to the SAGE house style: http://www.uk.sagepub.com/repository/binaries/pdf/SAGE_UK_style_guide_short.pdf"

Acknowledgements

We would like to thank everyone who supported us throughout our course of this master thesis.
This study was conducted under the lead of promoter Prof. Dr. P. Feys and co-promoter Dr. I. Baert. Their guidance was a great help. Therefore, we would like to express our gratitude towards them for their engagement and their support throughout our master thesis with their knowledge.
Furthermore, we would also like to thank all the participating persons with multiple sclerosis and the physical therapists at each clinical unit for their cooperation. Without their participation, this study would have been impossible. The assessment of all the measures was time-consuming for both the patients and the physical therapists. Accordingly, we are very grateful they were willing to spend their time for this research. We are also thankful that the physical therapists were willing to fill out all of the uniform data file templates with the results of the measures, as this was of great help for the data-collection and data-analysis.
Finally, we would like to thank the European Rehabilitation In Multiple Sclerosis (RIMS) Network and the Special Interest Group (SIG) on mobility which have initiated the development of multi-center studies. It was a great opportunity to be part of this study.

Leopoldsburg, 10-06-15 L.J.
Hasselt, 10-06-15 K.V.

Framework

This study is part of a duo master thesis of L. Jans and K. Vanoppen, students of rehabilitation sciences and physiotherapy, and contributes to the neurological rehabilitation.

This study mainly took place in the framework of the European Rehabilitation In Multiple Sclerosis (RIMS) Network for best practice and research (www.eurims.org), more specifically within the Special Interest Group (SIG) on Mobility. Additionally, some groups from the United States (US) participated.

The SIG mobility is a group of clinicians, mainly physiotherapists and rehabilitation doctors, as well as researchers with an interest in mobility-related topics in MS. The group gathers annually with the aim to exchange information; to strengthen networks between therapists, researchers and clinical centers and to implement and elaborate evidence-based rehabilitation.

Within the above context, the SIG Mobility Assembly has initiated the development of multi-center studies, as this allows the inclusion of a substantial higher number of participants compared to individual projects with a small sample size. It therefore can lead to robust findings and it allows subgroup analyses which is considered as crucial. From 2009, studies entitled MC-Walking Study part I ('Effect of time of the day on walking capacity'), MC-Walking Study part IIa ('Differential value of walking capacity tests') and MC-Walking Study part IIb ('Responsiveness of short and long walking capacity tests in multiple sclerosis') have been performed and resulted in several publications[1-4] and some more publications in the future.

This study is part of the third multi-center study (MCS-III-MOB) 'Psychometric properties of outcome measures of mobility beyond walking speed in multiple sclerosis' which is a sequel to previous work. This study concerns mobility beyond walking speed, as such not only walking measures (assessing walking speed) but also functional mobility measures were included (assessing balance and gait maneuvers used in everyday life).

In this part of the study, the convergent validity of mobility measures in pwMS was investigated. Drafting of the data file templates, assisting in part of the data-collection, the data quality revision, the data-analysis and the reporting of the results were executed by the students. The workload between L. J. and K. V. was equally distributed. In this study, only the data available on May first were included for investigating the convergent validity. However, data collection is still ongoing with more pwMS being assessed in the participating clinical units. Further assessment of the validity will be performed after completion of the data collection. Also the other psychometric properties (reliability and responsiveness) will be investigated in a next part of the study.

The presented study may provide insights in the suitability of mobility measures in persons with MS (pwMS). Psychometric evaluation of mobility measures in pwMS is important because outcome measures have to be valid, reliable and responsive for a population before any assumption can be

made based upon the patients' results. Mobility measures have been well investigated for psychometric properties in the stroke population, however this was limited investigated in the MS population.

1. Feys, P., et al., *Effect of time of day on walking capacity and self-reported fatigue in persons with multiple sclerosis: a multi-center trial.* Mult.Scler., 2012. **18**(3): p. 351-357.
2. Gijbels, D., B.O. Eijnde, and P. Feys, *Comparison of the 2- and 6-minute walk test in multiple sclerosis.* Mult.Scler., 2011. **17**(10): p. 1269-1272.
3. Gijbels, D., et al., *Which walking capacity tests to use in multiple sclerosis? A multicentre study providing the basis for a core set.* Mult.Scler., 2012. **18**(3): p. 364-371.
4. 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*. NeurorehabilNeural Repair. 2014.

Validity of mobility measures in multiple sclerosis, according disability level: a RIMS multi-center study

Laura Jans, PT, MSC[1], Kristien Vanoppen, PT, MSC[1]

Abstract

Background: Psychometric properties of walking measures in pwMS (persons with MS) are well investigated, however there is limited investigation of mobility measures.

Objective: To investigate convergent validity of mobility measures in pwMS, according disability level.

Methods: A non-controlled interventional multicenter study design was applied with a convenience sample (101 pwMS). There were seven clinician-reported and three patient-reported outcome measures (ClinROs and PROs) (101 pwMS), and five optional measures (48 pwMS). Disability subgroups were made according EDSS-level (Expanded Disability Status Scale). Spearman's rank and Pearson correlation coefficients were calculated.

Results: Overall, the highest validity was found for ClinROs. The highest, all significant, correlations were found for the Timed 25-Foot Walk (T25FW) (0.52-0.93), 2 Minutes Walking Test (2MWT) (0.54-0.93) and Performance Scale mobility (PSmob) (0.53-0.79) In the mild subgroup, the highest correlations were found. Low non-significant correlations were found in the moderate subgroup for the 5 Sit-To-Stand test (5STS) and modified 5STS (mod5STS), in the severe subgroup for the mod5STS and Multiple Sclerosis Walking Scale-12 (MSWS-12) and in the very severe subgroup for the PROs. The Timed Up and Go manual (TUG_{manual}) and Activities-specific Balance Confidence Scale (ABC) showed the highest validity in the moderate-severe subgroup of the optional measures.

Conclusion: For pwMS with a mild-moderate disability, a combination of ClinROs and PROs should be selected. For a (very) severe disability, a combination of ClinROs assessing different aspects of mobility should be used. To assess walking the T25FW and 2MWT are most valid, for balance and transitions, the TUG_{manual} and TUG are most valid.

Keywords

Multiple sclerosis, mobility, walking, balance, validity

[1]Hasselt University, REVAL Rehabilitation Research Center, Biomedical Research Institute (BIOMED), Diepenbeek, Belgium

Introduction

Walking difficulties are a common impairment in persons with multiple sclerosis (pwMS), with adverse effects on personal quality of life and quality of life of caregivers[1]. For 70% of pwMS difficulty in walking is the most challenging aspect of multiple sclerosis (MS). PwMS consider walking as the most valuable bodily function, with about 41% to 75% of people experiencing walking problems[1, 2]. Gait abnormalities already occur in an early stage in pwMS[3, 4]. Seventy four percent of pwMS experience that walking problems disrupt their daily life and only 34% of pwMS with walking difficulties are employed[1, 2].

Also balance deficits are common in pwMS[5, 6]. These deficits are characterized by an increased sway in stance, delayed responses to postural disturbances and a reduced ability to move towards their limits of stability[7]. PwMS fall frequently[7, 8], have a fear of falling[9] and their risk of fall-related injuries is increased[10]. This increased risk of falls has also been associated with the use of a walking aid[11].

Functional mobility is a term used to reflect the balance and gait maneuvers used in everyday life (e.g. walking, getting in and out of a chair and turning)[12]. In particular, mobility measures assess 'walking and moving' and 'changing and maintaining body positions (balance)' to allow the implementation of a more tailored treatment plan to improve mobility and reduce the risk of falling[13]. Mobility can be assessed by objective clinical scales as well as patient-reported questionnaires. In addition to scales measuring performance, the use of questionnaires measuring the patients perception and the behavioral consequences of mobility problems are important to assess the impact on activities of daily life[13].

Walking can be assessed by walking capacity measures. These encompass short-distance measures of walking speed (e.g. the 10 Meter Walk Test, 10mWT; Timed 25 Foot Walk, T25FW) as well as longer, timed measures of walking distance (e.g. the 6 Minute Walk Test, 6MWT; the 2 Minute Walk Test, 2MWT)[14]. Short walking tests are adequate for generally describing a MS patient's walking capacity[14]. For intervention studies, inclusion of a long walking distance test is advised[14]. To assess walking ability subjectively patient-rated outcomes measures such as the Multiple Sclerosis Walking Scale-12 (MSWS-12), are used[15, 16].

Mobility (including walking and balance) can be assessed by objective clinical scales. For example, the 5 Sit-To-Stand test (5STS), Berg Balance Scale (BBS), Four Square Step Test (FSST) and Trunk Impairment Scale (TIS) are scales which assess balance. And the Timed Up and Go (TUG) and Dynamic Gait Index (DGI) assess walking and balance.

Patient-reported questionnaires to assess mobility are the Rivermead Mobility Index (RMI), Dizziness Handicap Inventory (DHI), Activities-specific Balance Confidence Scale (ABC) and the Performance Scale mobility (PSmob). These measures include different component of mobility such as trunk movements, holding postures, reactions on internal and external disturbances, walking (backwards),

stepping, change of walking speed, change of walking direction, change of surface, stairs, turning and transfers.

Psychometric evaluation of these measures is important because outcome measures have to be valid (measuring the concept that it intends to measure), reliable (the measure is consistent and free from error) and responsive (the degree to which a measure detects change over time) for a population before any assumption can be made based upon the patients' results[17]. Mobility measures have been well investigated for psychometric properties in the stroke population, however this was limited investigated in the MS population.

Psychometric properties of walking capacity measures are well known in pwMS. According to a review by Kieseier et al., timed walking measures (T25FW, 10mWT, 2MWT and 6MWT) correlate strongly with each other and have a moderate to strong correlation with the MSWS-12[18]. Also responsiveness and clinical important change of these measures have been investigated[19]. Long walking measures (2MWT and 6MWT) and the MSWS-12 are more appropriate than short walking measures to detect clinically meaningful improvement after physical rehabilitation.

In our previous work we reviewed the psychometric properties of mobility measures in MS. Only the BBS and MSWS-12 have been investigated more thoroughly. Data concerning the reliability was found for 9 outcome measures (BBS, DGI, Kela Coordination Test, Melsbroek Disability Scoring Test (MDST), SSST, TUG, ABC, DHI and MSWS-12) and showed excellent reliability. The validity of 15 measures (SSST, TIS, 5STS-test, Balance Evaluation Systems Test (BESTest), DGI, MSWS-12, RMI, Activities and Participation Questionnaire (APQ), Romberg, BBS, Ambulation Index (AI), TUG, ABC, DHI, Short Form-36 Physical Inventory (SF-36PI)) showed a varying level of validity. Moreover, most outcome measures have only been investigated by one or two studies in small sample sizes, and there was no distinction between disability levels, indicating further research is necessary to confirm the psychometric properties. Also information about the responsiveness of the measurement tools was very limited.

This study will provide insights in the suitability of mobility measures in pwMS by investigating the convergent validity of these measures in pwMS, according disability level. Reliability and responsiveness will be discussed in another paper.

Methods

Participants

A convenience sample of 101 pwMS was recruited at 9 clinical units (table 1) within the European RIMS network for best practice and research in MS rehabilitation and US. Hobart et al. (2012) reported that sample sizes of a minimum of 80 for validity studies provided estimates highly representative of the main neurological study samples[20]. Included subjects had a definite diagnosis of MS according to the McDonald criteria[21] and preservation of at least some ambulatory function (Expanded Disability Status Scale [EDSS] ≥2 and ≤7 as determined by neurologists or trained clinicians)[22]. The subjects had not experienced a relapse within the last month and had no changes in disease modifying treatment and no corticoid-therapy within the last month. They receive at least 10 sessions of therapy (in- or outpatient rehabilitation), with a maximum duration of 3 months. Patients were excluded if they had other medical conditions interfering with mobility (e.g. pregnancy, fractures,...) or other neurological conditions with permanent damage (e.g. stroke, Parkinson,...). Also patients with MS like syndromes were excluded (e.g. neuromyelitis optica,...) and patients who were not able to understand and execute simple instructions.

Table 1. Participating clinical units.

No.	Country	Clinical unit	Number of patients
1	Belgium	National Multiple Sclerosis Center Melsbroek	10
2	Belgium	Rehabilitation and Multiple Sclerosis Center Overpelt	3
3	Belgium	De Mick, Brasschaat	1
4	Czech Republik	Charles University in Prague and General Faculty Hospital	26
5	Israel	Multiple Sclerosis Centre, Sheba Medical Center, Tel-Hashomer	3
6	Norway	Multiple Sclerosis Center Hakadal AS	24
7	Serbia	Clinical Center in Belgrado	6
8	Spain	Eugenia Epalza Rehabilitation Center Bilbao	17
9	United States	Shepherd Centre in Atlanta, GA	11
	Total		101

Experimental Design and Outcome Measures

Design A non-controlled interventional multicenter study design was applied. Patient characteristics measured at baseline were: (1) age, (2) gender, (3) Body Mass Index, (4) disease duration since diagnosis, (5) type of MS, (6) disability level (EDSS), (7) use of a walking device, (8) history of falling, (9) severity of fatigue (Fatigue Scale for Motor and Cognitive functions (FSMC)) and (10) cognitive function (Symbol Digit Modalities Test (SDMT)).

Mobility outcome was assessed pre and post rehabilitation in the same order including rest periods, according to a standardized test protocol, including details on test procedures, verbal instructions and level of encouragement. For the purpose of assessing validity, only the data from the assessment pre rehabilitation were used.

A summary of the study design and measures can be found in appendix 1.

Outcome measures To investigate validity of mobility in multiple sclerosis, several outcome measures were used. The selection was based on our previous literature review study, recommendations from the Multiple Sclerosis Outcome Measures Taskforce[23] and discussion rounds within the workgroup, besides feasibility and appropriateness for current research questions.

The clinician-reported outcome measures (ClinROs) were: (1) T25FW, (2) TUG, (3) TUG_{cogn}, (4) 2MWT, (5) FSST, (6) 5STS and (7) modified 5 Sit-To-Stand test (mod5STS).The patient-reported outcome measures (PROs) were: (1) RMI, (2) MSWS-12 and (3) PSmob.

These measures assess different constructs underlying functional mobility, in particular walking, balance and transitions. For each construct, at least two measures were selected. Also counter-matches between ClinROs and PROs were sought.

The T25FW and 2MWT were selected as ClinROs to assess walking.

The T25FW[24] is a short distance measure of walking speed. The subject is directed to one end of a clearly marked 25-foot course and is instructed to walk 25 feet as quickly as possible, but safely. Patients may use assistive devices when doing this task.

The 2MWT[25] is a long walking capacity test recording the maximal distance a subject walks at fastest speed possible in two minutes. The 2MWT appears more feasible and less burdensome for patients than the 6-minute walk test. The more distance covered, the better is the walking performance.

The TUG and TUG_{cogn} are ClinROs to assess walking, balance and transitions.

The TUG[12] is a simple performance-based measure of dynamic balance. The subject stands up from a chair, walks 3m, turns back, and sits down again as quickly and safely as possible while being timed.

The TUG_{cogn}[26] is adding a cognitive task (subtracting three from a random number between 20 and 100) while performing the TUG. Use of their customary walking aid is allowed.

The FSST[27] is a ClinRO used to assess dynamic standing balance. The person rapidly steps forward, sideways, backward and sideways over a low obstacle (e.g. a cane), first clockwise, then counterclockwise. The test is scored by the time it takes to perform the task. Shorter times indicate better balance.

The 5STS and mod5STS are ClinROs to assess balance and transitions. The 5STS-test[28] is a timed test where the subject is asked to stand up and sit down as quickly as possible five times from a standard chair with arms, keeping the arms folded across the chest. It is a performance-based multi-dimensional task that is a measure of both balance and lower extremity strength. We also created a modified version (mod5STS) in which persons are allowed to use their arms to push off from a chair with armrests.

The RMI is a PRO with a broad array of items which entail both constructs of walking, balance and transitions. The RMI[29] is an extension of the gross function subscale of the Rivermead Motor Assessment. It has a focus on disability, is simple and quick to administer, is able to be used in hospital and home settings and spans a wide range of reduction in mobility (turning over in bed to running). Higher scores indicate less mobility problems.

The MSWS-12[15] is a self-assessment scale which measures the impact of MS on walking. It consists of 12 questions concerning the limitations to walking due to MS during the past two weeks. Lower scores indicate less perceived walking impairment.

The PSmob assesses walking ability on an ordinally scored scale with zero being normal and six being total disability. It is a subscale of the PS[30], which uses a single question to assess each of eleven domains (mobility, hand function, vision, fatigue, cognition, bladder/bowel, sensory, spasticity, pain, depression and tremor/coordination).

Following measures were optional but strongly recommended: (1) TUG_{manual}, (2) BBS, (3) DGI, (4) ABC and (5) Trunk Impairment Scale modified Norwegian version (TIS-modNV).

The TUG_{manual} and ABC assess the functional mobility construct (balance and walking). The BBS assesses balance and the DGI assesses walking ability. The TIS-modNV assess a more specific aspect of balance, namely sitting balance.

The TUG_{manual} involves performing the TUG while holding a full cup of water.

The BBS[31] is a balance measure consisting of 14 functional balance tasks of increasing difficulty with scores ranging on an ordinal scale from zero to four, determined by the ability to perform the assessed activity. Two dimensions of balance are addressed: static balance (the ability to maintain upright posture) and dynamic balance (postural adjustments for voluntary movement). The maximum score is 56 points, with a higher score indicating better performance.

The DGI[32] assesses a subjects' ability to respond to changing task demands during walking. It consists of eight items including walking, walking with changing speeds, walking with head turns, pivoting, walking while stepping over or around obstacles and going up four stairs. Each item is scored on a four-point scale (0= severe impairment and 3= no gait dysfunction). The highest possible total score is 24 points.

The ABC[33] is a 16-item self-report questionnaire that asks people to rate their balance confidence in performing everyday activities on a numeric rating scale (range: 0-100). A score of zero represents no confidence and a score of 100 represents complete confidence in performing the activity.

The TIS-modNV[34] assesses static and dynamic sitting balance and trunk coordination in a sitting position. An ordinal scale is used to score the test items (0-3). The maximal score is 16, with a higher score indicating a better performance.

Statistical Analyses

Data-analysis was executed using SPSS. Subgroup stratification was done according to disability level, based on the EDSS-score. The sample was categorized based on the level of personal assistance and walking support: "mild" (EDSS 2-4), "moderate" (EDSS 4.5-5.5) requiring assistance or limitations of full daily activities, "severe" (EDSS 6) requiring unilateral support and "very severe" (EDSS ≥6.5) requiring bilateral support. Since there was more limited data availability for the optional outcome measures, these were only divided in two categories according to disability as either "mild" (EDSS 2-4) or "moderate-severe" (EDSS >4).

The mean and standard deviation of each continuous outcome measure and the median and interquartile range of each ordinal outcome measure was computed for the entire sample and according to EDSS-subgroups.

Normality of the data was checked using a Shapiro-Wilk test. Since the data were not normal distributed, a Kruskal-Wallis analysis was performed to investigate differences in mobility performance between subgroups.

Since there is no golden standard to assess mobility, concurrent validity could not be assessed. Therefore, convergent validity was evaluated by analyzing the degree to which each mobility measure correlates with the other mobility measures. For correlations of the entire sample, the non-parametric Spearman's rank correlation coefficients were used, since the data were non-normal distributed. When analyzing the data according to the subgroups, the parametric Pearson correlation coefficients were calculated for the normal distributed data. For the remainder of the non-normal distributed data, Spearman's rank correlations coefficients were used.

Convergent validity was analyzed according to EDSS-level to assess if validity varied according to disability level.

To summarize the evidence of convergent validity, a search for guidelines to interpret the results was conducted. Similar categories in different books and articles were found[35, 36]. Used benchmark values for interpreting results of psychometric properties are summarized in table 2.

We corrected for the number of tests by means of alfa adjustment (0.05/number of correlation tests).

Table 2. Benchmark values convergent validity.

Pearson and Spearman's rank correlation coefficients benchmark values	
|x| =0.00-0.24	Poor or no relationship
|x| =0.25-0.49	Fair degree of relationship
|x| =0.50-0.75	Moderate relationship
|x| >0.75	Good relationship
|x| >0.9	Excellent relationship

Abbreviations: |x|: true value of the correlation coefficient

Results

Patient characteristics and outcome measures

Patient characteristics are summarized in table 3.

An overview of the outcome measures is presented in table 4.

For the T25FW, TUG, TUG_{cogn}, 2MWT and FSST, differences between subgroups were all statistically significant, except when comparing the performance of EDSS 4.5-5.5 with EDSS 6.

For the 5STS, there was a significant difference between EDSS 2-4 and 6, and between EDSS 2-4 and EDSS ≥6.5, all other comparisons were not significant. Performance of the mod5STS in the very severe disability group (EDSS ≥6.5) was statistically significant worse than performance in all other subgroups, the differences between the other three subgroups mutually were not significant. For the RMI, there were only significant differences between EDSS ≤4 and all other categories. The remainder of comparisons were not significant.

For the MSWS-12, there was a significant difference between EDSS 2-4 and EDSS 4.5-5.5, and between EDSS 2-4 and EDSS≥6.5. Since there was a large interquartile range for EDSS 6 (45.58-89.85), there were no significant differences between EDSS 6 and the other categories. There was no significant difference between EDSS 4.5-5.5 and EDSS≥6.5.

For the PSmob differences between all subgroups were significant, except for the differences between EDSS 4.5-5.5 and EDSS 6.

For all optional outcome measures except for the ABC, a statistically significant difference was found between the performance of the mild (EDSS 2-4) and moderate-severe (EDSS >4) subgroup.

Table 3. Patient characteristics (n=101).

Age, years (mean±SD)	48.51±10.22
Gender: M/F (n,%)	45/56 (44.6/55.4)
BMI (mean±SD)	24.16±4.41
Employment status (n,%)	
Unemployed/disability	49 (48.5)
Working (full time/part time)	29 (28.7)
Retired	22 (21.8)
Unknown	1 (1)
Living arrangement (n,%)	
Alone	17 (16.8)
With family/partner	83 (82.2)
Unknown	1 (1)
Years since diagnosis (mean±SD)	10.97±8.17
Type of MS (n,%)	
PP	21 (20.8)
SP	25 (24.8)
RR	50 (49.5)
Unknown	5 (5)
EDSS (median, IQr)	5.50 (4.25-6.75)
EDSS 2-4 (n,%)	37 (36.6)
EDSS 4.5-5.5 (n,%)	23 (22.8)
EDSS 6 (n,%)	18 (17.8)
EDSS ≥ 6.5 (n,%)	21 (20.8)
Unknown (n,%)	2 (2)
Walking aid (n,%)	
No walking aid	48 (47.5)
1 crutch/cane	24 (23.8)
2 crutches/canes	14 (13.9)
Rollator	12 (11.9)
Wheelchair	2 (2)
Unknown	1 (1)
Orthosis (n,%)	
Yes	87 (86.1)
No	14 (13.9)
Number of falls in past 6 months (mean±SD)	3.14±8.32
Injurous falls in past 6 months (mean±SD)	0.29±1.21
SDMT (median, IQr)	40.00 (31.5-48.5)
FSMC (median, IQr)	
FSMCcogn	29.00 (19.5-38.5)
FSMCmot	36.00 (30.5-41.5)
FSMCtotal	66.00 (52-80)

Abbreviations: BMI: Body Mass Index; EDSS: Expanded Disability Status Scale; F: Female; FSMCcogn: Fatigue Scale for Motor and Cognitive functions, cognitive; FSMCmot: Fatigue Scale for Motor and Cognitive functions, motor; IQr: Interquartile range; M: Male; n: number of patients; PP: Primary Progressive; RR: Relapsing Remitting; SDMT: Symbol Digit Modality Test; SD: Standard Deviation; SP: Secundary Progressive

Table 4. Overview of the mobility outcome measures (mean±standard deviation; median (IQR)).

Outcome measures					
Outcome measure	**Total sample (n=101)**	**EDSS 2-4 (n=37)**	**EDSS 4.5-5.5 (n=23)**	**EDSS 6 (n=18)**	**EDSS ≥ 6.5 (n=21)**
T25FW (s)	9.05±6.69	5.39±1.78 [2,3,4]	8.40±2.91 [1,4]	8.49±2.83 [1,4]	17.93±11.21 [1,2,3]
TUG (s)	12.90±7.68	7.61±2.40 [2,3,4]	12.06±3.70 [1,4]	13.29±3.53 [1,4]	23.86±9.13 [1,2,3]
TUGcogn (s)	14.16±8.09	8.88±3.61 [2,3,4]	13.05±3.77 [1,4]	14.20±4.23 [1,4]	25.87±8.84 [1,2,3]
2MWT (m)	121.59±59.20	174.70±44.77 [2,3,4]	104.30±32.66 [1,4]	94.91±29.37 [1,4]	57.41±26.89 [1,2,3]
FSST (s)	15.05±10.31	9.30±3.85 [2,3,4]	14.83±4.66 [1,4]	19.75±11.81 [1,4]	33.93±13.61 [1,2,3]
5STS (s)	15.96±6.98	12.77±5.00 [3,4]	16.15±5.27	18.41±6.56 [1]	24.15±9.35 [1]
mod5STS (s)	14.43±5.91	11.80±4.18 [4]	12.98±3.17 [4]	14.85±4.32 [4]	21.23±7.43 [1,2,3]
RMI*	13.00 (12.00-14.00)	14.00 (13.50-14.50) [2,3,4]	13.00 (12.00-14.00) [1]	11.50 (10.00-13.00) [1]	12.00 (11.00-13.00) [1]
MSWS-12*	58.33 (38.02-78.65)	41.67 (22.40-60.94) [2,4]	64.58 (48.96-80.21) [1]	67.71 (45.58-89.85)	77.08 (62.50-91.66) [1]
PSmob*	3.00 (2.00-4.00)	1.00 (0.00-3.00) [2,3,4]	3.00 (3.00-3.00) [1,4]	3.00 (2.50-3.50) [1,4]	5.00 (4.50-5.50) [1,2,3]
Optional outcome measures					
Outcome measure	**Total sample (n=48)**	**EDSS 2-4 (n=15)**		**EDSS >4 (n=33)**	
TUGmanual (s)	13.83±8.20	8.54±2.65 [5]		16.91±8.79 [1]	
BBS*	43.50 (32.00-55.00)	53.00 (48.00-58.00) [5]		38.00 (26.50-49.50) [1]	
DGI*	15.00 (11.00-19.00)	21.50 (17.50-25.50) [5]		14.00 (10.50-17.50) [1]	
TIS-modNV*	9.00 (7.00-11.00)	10.00 (7.50-12.50) [5]		7.00 (3.50-10.50) [1]	
ABC	54.37±19.20	65.18±19.20		49.78±17.53	

* : median (IQR)

[1] : significant difference with EDSS ≤4; [2] : significant difference with EDSS 4.5-5.5; [3] : significant difference with EDSS 6; [4] : significant difference with EDSS ≥6.5; [5] : significant difference with EDSS >4

Abbreviations: 2MWT: 2 Minutes Walking Test; 5STS: 5 Sit-To-Stand; ABC: Activities-specific Balance Scale; BBS: Berg Balance Scale; DGI: Dynamic Gait Index; EDSS: Expanded Disability Status Scale; FSST: Four Square Step Test; IQr: Interquartile Range; mod5STS: modified 5 Sit-To-Stand test; m: meter; MSWS-12: Multiple Sclerosis Walking Scale-12; n: number of patients; PSmob: Performance Scale mobility; RMI: Rivermead Mobility Index; s: seconds; T25FW: Timed 25 Feet Walking Test; TIS-modNV: Trunk Impairment Scale-modified version; TUG: Timed Up and Go; TUGcogn: Timed Up and Go cognitive; TUGmanual: Timed Up and Go manual

Convergent validity

Figure 1 gives an overview of the convergent validity of the outcome measures. An overview of the correlations of the mobility outcome measures for the entire sample can be found in table 5. After correcting for multiple testing, most correlations were still significant.

When the correlations with all the other measures were considered, the highest correlations were found for the T25FW (0.52-0.93) and 2MWT (0.54-0.93), showing moderate to excellent overall validity and for the TUG (0.49-0.94) and the TUG_{cogn} (0.44-0.94), showing fair to excellent validity. The FSST, 5STS and mod5STS have fair to good validity (resp. 0.41-0.88, 0.41-0.88 and 0.38-0.88). The 5STS showed all significant correlations except for two non-significant correlations with the TUG_{manual} and the TIS-modNV. Also the mod5STS showed one non-significant correlation with the TUG_{manual}. The TUG_{manual} has fair to excellent overall validity(0.38-0.97). The BBS and DGI have fair to good validity (resp. 0.47-0.87 and 0.44-0.85). The TIS-modNV has four non-significant correlations and has shown to have fair to good validity (0.42 to 0.74).

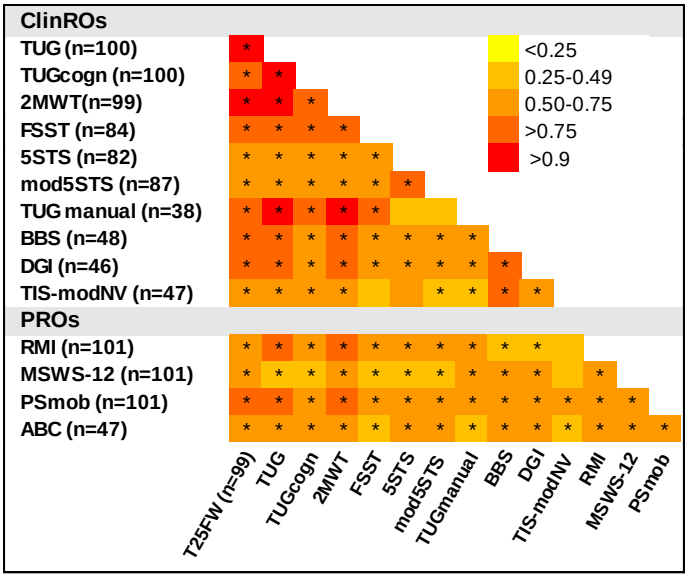


Figure 1. Convergent validity of the outcome measures.

Note: Non-parametric Spearman's rank correlation coefficients

Abbreviations: *: p<0.05/number of tests; 2MWT: 2 Minutes Walking Test; 5STS: 5 Sit-To-Stand; ABC: Activities-specific Balance Scale; BBS: Berg Balance Scale; ClinROs: clinician-reported outcome measures; DGI: Dynamic Gait Index; EDSS: Expanded Disability Status Scale; FSST: Four Square Step Test; mod5STS: modified 5 Sit-To-Stand test; MSWS-12: Multiple Sclerosis Walking Scale-12; n: number of patients; PROs: patient-reported outcome measures; PSmob: Performance Scale mobility; RMI: Rivermead Mobility Index; T25FW: Timed 25 Feet Walking Test; TISmodNV: Trunk Impairment Scale-modified version; TUG: Timed Up and Go; TUGcogn: Timed Up and Go cognitive; TUGmanual: Timed Up and Go manual

For the PROs, the RMI has fair to good validity (0.44-0.77). The MSWS-12 and the ABC have shown to have fair to moderate validity (resp. 0.41-0.63 and 0.44-0.69) and the PSmob moderate to good validity (0.53-0.79). Of the PROs, the RMI has the highest correlations with the ClinROs (0.44-0.77), showing fair to good validity.

When the correlations between mutual ClinROs were considered, the highest correlations were found for four timed measures: the T25FW, TUG, TUG_{cogn} and 2MWT, all showing moderate to excellent validity. When the correlations between mutual PROs were considered, a moderate validity was found for all of them.

As such, the overall validity of the ClinROs ranged from fair to excellent, while the overall validity of the PROs only ranged from fair to good. The validity of the mutual PROs was lower than the validity of the mutual ClinROs. The overall validity of the ClinROs and the validity of mutual ClinROs were both

shown to be fair to excellent. The overall validity of the PROs showed a more varying range of validity (fair to good) than the validity of mutual PROs (moderate).

Table 5. Correlations of the mobility outcome measures.

	T25FW	TUG	TUGcogn	2MWT	FSST	5STS	mod5STS	TUG manual	BBS	DGI	TIS-mod	RMI	MSWS-12	Psmob	ABC
ClinROs															
T25FW (n=99)	1.00														
TUG (n=100)	0.92*	1.00													
TUGcogn (n=100)	0.84*	0.94*	1.00												
2MWT(n=99)	-0.93*	-0.92*	-0.84*	1.00											
FSST (n=84)	0.86*	0.88*	0.78*	-0.85*	1.00										
5STS (n=82)	0.67*	0.71*	0.67*	-0.65*	0.63*	1.00									
mod5STS (n=87)	0.65*	0.66*	0.62*	-0.62*	0.55*	0.88*	1.00								
TUG manual (n=38)	0.84*	0.97*	0.85*	-0.92*	0.83*	0.43	0.38	1.00							
BBS (n=48)	-0.87*	-0.76*	-0.73*	0.80*	-0.59*	-0.59*	-0.71*	-0.68*	1.00						
DGI (n=46)	-0.78*	-0.76*	-0.73*	0.77*	-0.67*	-0.51*	-0.56*	-0.72*	0.85*	1.00					
TIS-modNV (n=47)	-0.70*	-0.61*	-0.56*	0.68*	-0.46	-0.50	-0.47*	-0.49*	0.77*	0.75*	1.00				
PROs															
RMI (n=101)	-0.72*	-0.77*	-0.71*	0.76*	-0.71*	-0.55*	-0.56*	-0.61*	0.47*	0.44*	0.44	1.00			
MSWS-12 (n=101)	0.52*	0.49*	0.44*	-0.54*	0.41*	0.41*	0.45*	0.51*	-0.56*	-0.50*	-0.42	-0.50*	1.00		
PSmob (n=101)	0.78*	0.78*	0.74*	-0.79*	0.70*	0.56*	0.54*	0.75*	-0.75*	-0.67*	-0.55*	-0.75*	0.63*	1.00	
ABC (n=47)	-0.65*	-0.59*	-0.54*	0.60*	-0.46	-0.51*	-0.69*	-0.42*	0.67*	0.58*	0.47*	0.50*	-0.56*	-0.53*	1.00

Note: Non-parametric Spearman's rank correlation coefficients

Abbreviations: *: significant at p<0.05/number of tests; 2MWT: 2 Minutes Walking Test; 5STS: 5 Sit-To-Stand; ABC: Activities-specific Balance Scale; BBS: Berg Balance Scale; ClinROs: clinician-reported outcome measures; DGI: Dynamic Gait Index; FSST: Four Square Step Test; mod5STS: modified 5 Sit-To-Stand test; MSWS-12: Multiple Sclerosis Walking Scale-12; n: number of patients; PROs: patient-reported outcome measures; PSmob: Performance Scale mobility; RMI: Rivermead Mobility Index; T25FW: Timed 25 Feet Walking Test; TIS-modNV: Trunk Impairment Scale-modified version; TUG: Timed Up and Go; TUGcogn: Timed Up and Go cognitive; TUGmanual: Timed Up and Go manual

Convergent validity of primary outcome measures according to disability

Overall, clinician-reported mobility measures showed high correlations with each other across all EDSS-levels, with lowest correlations with the TUG_{cogn}. The correlations between the mutual ClinROs were generally higher than the correlations between the ClinROs and the PROs, and the correlations between mutual PROs. An overview of the convergent validity of the primary outcome measures according to disability can be found in figure 2. For more detailed information about the correlations, overview tables are provided in appendix 2.

Mild disability (EDSS 2-4) For the primary outcome measures, the highest correlations were found in the mild disability group.

Overall, the most valid ClinROs were the 2MWT and the T25FW. The 2MWThas shown to have moderate to excellent validity and the T25FWmoderate to good validity. The TUG and FSST showed the greatest range of validity from fair to good. Also the PROs showed a great range of validity. The correlations of the RMI showed fair to moderate validity. The PSmob and MSWS-12 have shown to have fair to good validity.

The highest correlations were again found between mutual ClinROs, especially for theT25FW and the 2MWT, which have shown to have moderate to good validity. The correlations between mutual PROs were also high, with a moderate to good validity of the MSWS-12 and PSmob, and a moderate validity of the RMI.

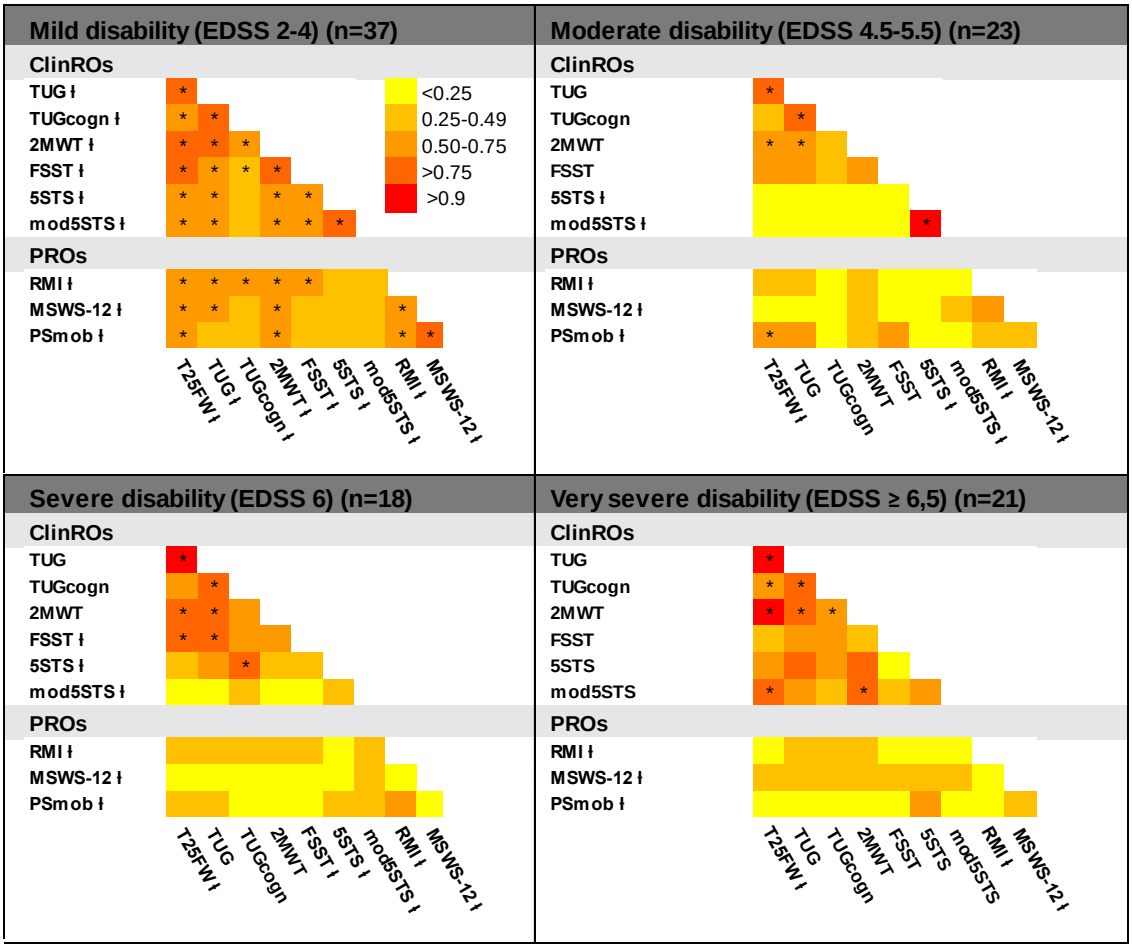


Figure 2. Convergent validity of primary outcome measures according to EDSS-level.
Abbreviations: *: p<0.05/number of tests; ł: non-parametric Spearman's rho correlation; 2MWT: 2 Minutes Walking Test; 5STS: 5 Sit-To-Stand; ClinROs: clinician-reported outcome measures; EDSS: Expanded Disability Status Scale; FSST: Four Square Step Test; mod5STS: modified 5 Sit-To-Stand test; MSWS-12: Multiple Sclerosis Walking Scale-12; n: number of patients; PROs: patient-reported outcome measures; PSmob: Performance Scale mobility; RMI: Rivermead Mobility Index; T25FW: Timed 25 Feet Walking Test; TUG: Timed Up and Go; TUGcogn: Timed Up and Go cognitive

Moderate disability (EDSS 4.5-5.5) For the moderate disability group, a few significant high correlations were found. Both ClinROs and PROs showed a varying level of overall validity. For the ClinROs, the T25FW and TUG are the most valid, showing poor to good validity except for low correlations with the 5STS and mod5STS. However, poor validity was found for the 5STS and poor to fair validity for the mod5STS. The only exception was the correlation of these two measures with each other, which was excellent. For the PROs, the MSWS-12 has shown to have only poor to fair validity, except for a moderate correlation with the RMI.

The highest correlations were again found between mutual ClinROs, with the highest correlations for the TUG, except for the low correlations of this measure with the 5STS and mod5STS.

The correlations between mutual PROs were quite low with a fair to moderate validity for the RMI and MSWS-12 and a fair validity for the PSmob.

Severe disability (EDSS 6) For pwMS with severe disability, again a few significant high correlations were found. Both ClinROs and PROs showed a very varying level of overall validity. For the ClinROs, the highest correlations were found for the T25FW and the TUG, with validity ranging from poor to excellent. All the ClinROs showed the lowest correlations with the PROs and also a very low correlations with the mod5STS. The mod5STS has shown to have poor to fair validity. The PROs showed low overall validity, with a poor to fair validity of the MSWS-12 and a poor to moderate validity of the PSmob and the RMI.
The highest correlations were again found between mutual ClinROs. The TUG showed the highest validity (moderate to excellent), followed by the TUG_{cogn} (moderate to good), except for low correlations of these two measures with the mod5STS. The mod5STS has showed to have the lowest correlations with all the other ClinROs, showing poor to fair validity.
The correlations between mutual PROs were quite low, with a poor to moderate validity for the RMI and PSmob and a poor validity for the MSWS-12.

Very severe disability (EDSS ≥6.5) Also in the very severe disability category, a few significant high correlations were found and both ClinROs and PROs showed a very varying level of overall validity. The ClinROs showed the lowest correlations with the PROs. The FSST showed the poorest overall validity of all the ClinROs (poor to moderate). This measure also had low correlations with the other ClinROs, showing poor to moderate validity.
The PROs showed overall poor to fair validity, except for one high correlation between the PSmob and the 5STS.
Again, the ClinROs showed the highest correlations with each other. The T25FW, 2MWT and TUG have the highest correlations, showing moderate to excellent validity, except for the low correlations with the FSST. The correlations between mutual PROs showed a low level of validity, with the RMI showing poor validity and the MSWS-12and PSmob showing poor to fair validity.

Convergent validity of optional outcome measures according to disability

Correlation coefficients for the optional outcome measures (ClinROs: TUG_{manual}, BBS, DGI, TIS-modNV; PRO: ABC) according to two disability levels are presented in figure 3 and can be found in appendix 3.

Mild disability (EDSS 2-4) For the ClinROs, the TUG_{manual} and BBS have shown to have the highest overall validity in the mild disability subgroup.

The TUG_{manual} has shown to have fair to excellent validity, with the highest correlations with the TUG, TUG_{cogn} and 2MWT and also a very high correlation with the PSmob. The BBS and DGI have shown to have poor to good validity. The TIS-modNV had no significant correlations. It showed poor to moderate validity with very low correlations with the ABC and the mod5STS.

The only PRO, the ABC, showed overall fair to good validity, except for one very low correlation with the TIS-modNV.

When the correlations between mutual ClinROs were considered, the same level of validity was found as the overall validity.

When the correlations between the ABC and mutual PROs were considered, it showed moderate to good validity.

Moderate-severe disability (EDSS>4) In the moderate-severe disability group, the TUG_{manual}, which is one of the ClinROs, has shown to have the highest overall validity (fair to excellent) and six significant correlations were found. The highest correlations were found with the TUG and 2MWT.

The BBS and DGI both showed to have poor to good validity.

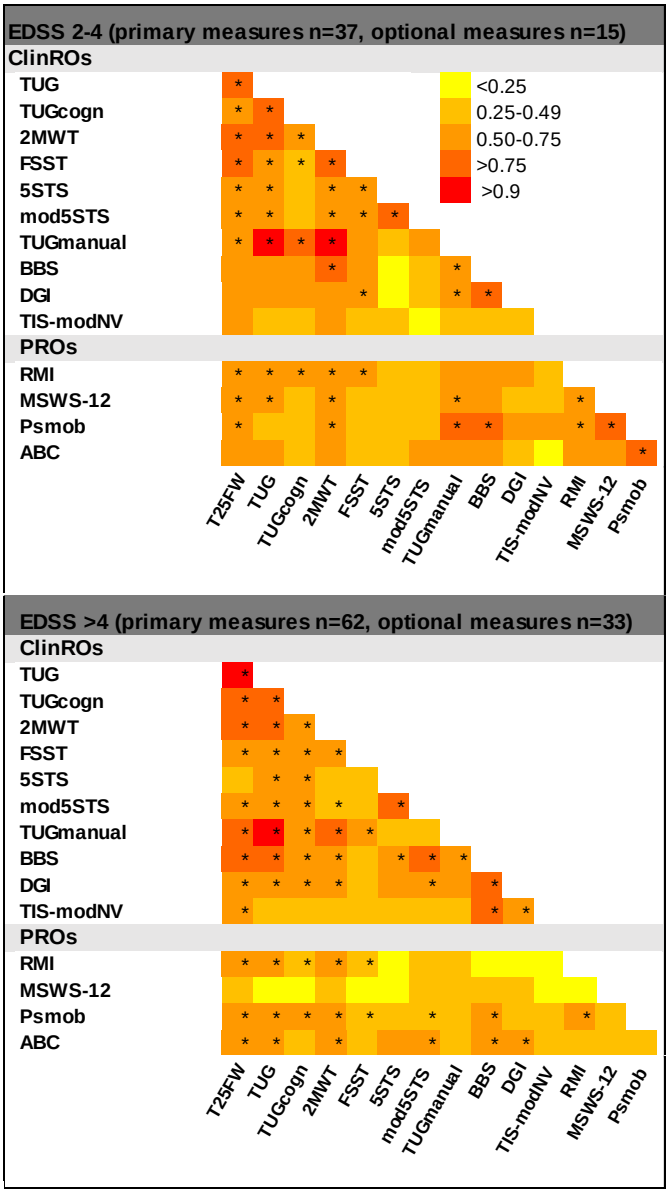


Figure 3. Convergent validity of outcome measures according to disability level (EDSS).

Note: Non-parametric Spearman's rank correlation coefficients

Abbreviations: *: p<0.05/number of tests; 2MWT: 2 Minutes Walking Test; 5STS: 5 Sit-To-Stand; ABC: Activities-specific Balance Scale; BBS: Berg Balance Scale; ClinROs: clinician-reported outcome measures; DGI: Dynamic Gait Index; EDSS: Expanded Disability Status Scale; FSST: Four Square Step Test; mod5STS: modified 5 Sit-To-Stand test; MSWS-12: Multiple Sclerosis Walking Scale-12; n: number of patients; PROs: patient-reported outcome measures; PSmob: Performance Scale mobility; RMI: Rivermead Mobility Index; T25FW: Timed 25 Feet Walking Test; TISmodNV: Trunk Impairment Scale-modified version; TUG: Timed Up and Go; TUGcogn: Timed Up and Go cognitive; TUGmanual: Timed Up and Go manual

The TIS-modNV has only three significant correlations and showed fair to good validity. The highest correlations were found with the BBS and DGI.

The ClinROs showed the highest correlations with mutual ClinROs. The only PRO, the ABC has shown to have fair to moderate overall validity. When the correlations between the ABC and mutual PROs were considered, it showed lower correlations, indicating a fair validity.

Overview of convergent validity

An overview of the convergent validity of the outcome measures for the entire sample and according to EDSS-level can be found in appendix 3.

Discussion

The present study gives insight in the suitability of mobility measures in pwMS by assessing the convergent validity of these measures in pwMS, according disability level. Generally, the overall validity of the ClinROs (fair to excellent) was higher than the overall validity of the PROs (fair to good). Low correlations between mutual PROs were found, especially for pwMS with more disability. The timed walking measures showed a high overall validity, including the different versions of the TUG. Except for the TUG_{cogn}, which showed very few significant correlations. The TIS-modNV showed four non-significant correlations when looking at the overall validity. The mod5STS had the highest correlation with the 5STS in all the subgroups when the correlations with the ClinROs were considered, except in the very severe disability group (EDSS ≥6.5). The FSST showed poorest validity of all the ClinROs for pwMS with very severe disability. The DGI showed three significant correlations in EDSS 2-4 compared to eight in EDSS ≥4. Low correlations of the PROs were especially found in the higher disability subgroups (EDSS 6 and ≥6.5). Moreover, the correlations between the mutual PROs (moderate) were lower than between the mutual ClinROs (fair to excellent). This difference was most prominent in the subgroup of pwMS with more disability (EDSS >4). This could indicate that the ClinROs are more valid to assess mobility in pwMS. The low mutual correlations of the PROs could also indicate that these measures assess different constructs of mobility, and are as such complementary in the information they provide.

For the walking measures, high validity was found in all the subgroups for the T25FW, 2MWT and the TUG. These measures all assess walking speed. Since their construct is very similar, correlations between these measures were high.

Correlations between the different versions of the TUG, incorporating either motor-motor (TUG_{manual}) or motor-cognitive (TUG_{cogn}) dual tasking were good to excellent. Also the correlations between the TUG_{manual} and the T25FW and 2MWT were good to excellent, indicating these measures have a similar construct. This indicates that the TUG_{manual} is valid to assess mobility in pwMS. In all disability subgroups, very few significant correlations were found for the TUG_{cogn}. This could be because the TUG_{cogn} involves the addition of a cognitive task and as such evaluates cognition rather than mobility. Cognitive impairments are often present in pwMS (up to 70%) and these impairments occur independently of the physical disability assessed by the EDSS-level[37]. The SDMT has been assessed to evaluate the cognitive function, in future research a regression analysis can be executed to investigate if the SDMT is a predictor for the outcome of the TUG_{cogn}.

When comparing our results with previous research, mostly similar results were found. A review by Kieseier et al. showed fair to good validity of the T25FW[18]. Three correlations with the MSWS-12 were found (0.46-0.78) and also correlations with three other measures (TUG: 0.85, DGI: 0.59, PSmob: 0.77). All these correlations, except for the correlation with the DGI, are in line with our results. With the DGI, we found a much higher correlation (0.78). The lower correlation in previous research could be due to the very small sample size (n=21)[38]. The correlations of the TUG with three other measures

were previously evaluated (BBS: 0.62, DGI: 0.72, ABC: 0.30)[39]. Correlations with the BBS and DGI were consistent with our results, however a higher correlation with the ABC was found (0.59). Gijbels et al. found a correlation of 0.72 between the 2MWT and the MSWS-12[40]. We found a lower correlation of 0.54. This could be due to the higher median EDSS-score of our sample (5.50), compared to their sample (3.50). In previous research, the 2MWT was shown to be the most responsive walking measure. Since validity of the T25FW, 2MWT and TUG is similar, but the responsiveness differs highly, the 2MWT is most recommended to evaluate walking[19].

For the overall validity, there were four non-significant correlations with the TIS-modNV (FSST, 5STS, RMI and MSWS-12). This could be explained by the fact that this measure assesses trunk mobility during sitting, while most other measures comprise coming from sit to stand or walking. However, this measure can be useful to include in the core mobility assessment of pwMS. A significant difference in the performance of the TIS-modNV in pwMS with EDSS ≤4 and >4 was found, which indicates that this measure can be valid to assess change in trunk mobility. To evaluate the validity of the TIS, only correlations with the FIM and EDSS have been evaluated previously[41], as such no comparison with previous research could be done.

When the correlations with the ClinROs were considered, the mod5STS had the highest correlation with the 5STS in all the subgroups except in the very severe disability group (EDSS ≥6.5). In this subgroup, the mod5STS has shown to have a lower correlation with the 5STS and a poor to good validity. Also, 11 of the 21 patients in this subgroup were unable to perform the 5STS. This could indicate that the mod5STS is better suited for this subgroup, because it is an easier version than the 5STS and is as such more easily performed by persons with higher disability levels. The mod5STS only showed a significant difference between EDSS ≥6.5 and the other subgroups. No statistical significant differences between other disability groups were found. As such, the mod5STS has poor discriminant validity in pwMS with an EDSS ≤6. In these pwMS, the 5STS is more valid. In previous research, validity of the 5STS has only been evaluated by correlations with muscle strength in the most affected leg and body sway performance with open eyes[42], as such this can not be compared to our results.

For pwMS with very severe disability, the FSST had the lowest overall validity of all the ClinROs (poor to moderate). It showed only poor to fair validity considering the correlations with the other ClinROs, except for one moderate correlation with the TUG. Also, only eight of the 21 patients in this subgroup could perform this measure. As such, this measure could be inappropriate to evaluate mobility in this subgroup because it may be too difficult to perform for pwMS with very severe disability. The FSST also showed poor correlations with the other ClinROs assessing balance (TUG_{cogn} and mod5STS), which could be explained by the fact that the FSST also assesses coordination besides balance.

The BBS showed overall higher correlations with the timed walking measures (T25FW, TUG, 2MWT) compared to the correlations with the FSST, 5STS and mod5STS, which also assess balance. This

could be because the FSST also assesses coordination as previously mentioned and the 5STS and mod5STS also assess lower extremity strength. In previous research, correlations of the BBS were fair to good (TUG: 0.62, DGI: 0.78, ABC: 0.48)[39]. We found slightly higher correlations with these measures (TUG: 0.76, DGI: 0.85, ABC: 0.67). The lower correlations in previous research could be due to a smaller sample size (n=51). Also the EDSS-scores of the patients in this research are unknown.

The validity of the DGI in both subgroups of pwMS was poor to good, but it only showed three significant correlations in EDSS 2-4 compared to eight in EDSS ≥4. This could indicate that the DGI is not relevant to assess mobility in pwMS with less disability but it is particularly relevant in pwMS with higher disability, because of the risk of falling and the prevalence of imbalance. Previous research found moderate to good validity for the DGI (T25FW: 0.78, TUG: 0.76, FSST: 0.77, BBS: 0.78, MSWS-12: 0.72, ABC: 0.54)[39, 43]. Similar correlations were found, except for the correlation with the MSWS-12, which was lower (0.50).

In general, the overall validity of the PROs was fair to good. Low correlations were especially found in the higher disability subgroups (EDSS 6 and ≥6.5), poor validity of the PROs in these subgroups could be due to a ceiling effect of these questionnaires. However, no ceiling effects of the PROs were found. The PROs also showed low mutual correlations, which could indicate that these measures assess different constructs of mobility, as previously mentioned. In previous research of the MSWS-12, correlations with the T25FW[44] and PSmob[16] were found (resp. 0.65 and 0.77). We found considerably lower correlations (resp. 0.52 and 0.63). In previous research of the RMI, its correlation (0.49) with habitual walking performance was considered (the real amount of steps performed in the customary living environment). As such, this cannot be compared to our results[45]. Previously, the ABC has been shown to have a fair to moderate validity (TUG: 0.38, BBS: 0.48, DGI: 0.54)[39]. We found higher correlations with the TUG and BBS (resp. 0.59 and 0.67). And finally, in previous research of the PSmob, two correlations were found (T25FW: 0.77, MSWS-12: 0.77)[16, 18]. We found similar correlations with the T25FW, but the correlation with the MSWS-12 was lower (0.63). For the overall validity of the PSmob, moderate to good valdity was found in this study. When assessing the validity according to EDSS-level however, validity was only fair to good for EDSS 2-4 and poor to moderate for EDSS>4. This could be due to data distribution, which was more normal distributed for the entire sample. Also sensitivity to change is questionable, since it is only a seven-point scale. The difference in validity of the PROs with respect to previous research could be due to a different disability level of the sample. Very often only the mean or median EDSS-score or the mean years since the onset of the pathology were documented. Also the type of MS and the age of the sample could have influenced the results.

When looking at the statistically significant differences in the results of the outcome measures between subgroups, there is generally a progressive worsening of performance according to a higher disability level. As such, mobility performance worsened with higher EDSS-level, a logical conclusion

considering higher EDSS-level indicates more disability. Most differences between disability subgroups were significant. This indicates that these measures could possibly discriminate between pwMS with different disability levels based on the EDSS-score. The T25FW, TUG, TUG_{cogn}, 2MWT and FSST showed all significant differences between subgroups, except when comparing the performance of EDSS 4.5-5.5 with EDSS 6. As such, these mobility measures were not able to discriminate between pwMS with moderate or severe disability.

Finally, pwMS with lower disability levels showed more significant correlations, especially for the objective measures. It is possible that some objective measures may be too difficult for patients with more disability or too easy for patients with less disability. With our sample however, no floor or ceiling effects were found. It is also possible that in pwMS with a higher disability the different constructs cannot be assessed with one outcome measure, but a combination of outcome measures assessing different aspects of functional mobility (balance, walking, transitions) should be selected.

All the mobility measures have a varying level of validity, ranging from fair to excellent. This emphasizes that diverse scales may assess different components of functional mobility and no single measure is appropriate to measure mobility of every pwMS. However, not only the validity of mobility measures should be considered, also reliability, responsiveness and clinical utility determine the suitability of mobility measures in pwMS.

Implications for Practice and Research

Practice

The T25FW and 2MWT are the most valid ClinROs to assess mobility in pwMS.

For the PROs, a combination of questionnaires is advised, as these seem to be complementary in content.

Since there is a great range of validity of the ClinROs and PROs for pwMS with mild (EDSS 2-4) or moderate (EDSS 4.5-5.5) disability, the best method to evaluate mobility in these patients is to select a combination of ClinROs and PROs to get a more clear picture of the mobility problems.

For pwMS with severe (EDSS 6) or very severe (EDSS ≥6.5) disability, capacity measures are more valid. Also a combination of outcome measures assessing different aspects of functional mobility (balance, walking, transitions) is necessary.

The FSST could be too difficult to perform in pwMS with an EDSS ≥6.5, other ClinROs should be selected.

The DGI should be used to assess walking in pwMS with higher disability (EDSS ≥4) because of the risk of falling and the prevalence of imbalance.

An individual selection of mobility measures should be done. Hereby keeping the disability level in mind, since the appropriateness of the mobility measures differs according to EDSS-level.

Recommendations to evaluate mobility in the disability subgroups of pwMS according to constructs (walking, balance and transitions) can be found in table 6. Measures are listed according to their validity, starting with the most valid measure. Some mobility measures, such as the TUG and RMI, assess all different aspects of mobility.

Table 6. Recommendations of the most valid mobility measures in pwMS, according to constructs.

		EDSS 2-4	**EDSS 4.5-5.5**	**EDSS 6**	**EDSS ≥ 6.5**
Walking	**ClinROs**	2MWT T25FW	T25FW TUG TUGcogn DGI	T25FW TUG	T25FW TUG 2MWT
	PROs	PSmob MSWS-12	PSmob RMI MSWS-12	RMI PSmob	PSmob
Balance	**ClinROs**	TUGmanual TUG TUGcogn FSST	TUGmanual BBS TIS-modNV	TUGmanual TUG TIS-modNV	TUGmanual TUG TIS-modNV
	PROs	ABC RMI	ABC RMI	ABC RMI	ABC RMI
Transitions	**ClinROs**	TUGmanual TUG TUGcogn	TUGmanual TUG TUGcogn	TUGmanual TUG	TUGmanual TUG
	PROs	RMI	RMI	RMI	RMI

Abbreviations: 2MWT: 2 Minutes Walking Test; ABC: Activities-specific Balance Scale; BBS: Berg Balance Scale; ClinROs: clinician-reported outcome measures; DGI: Dynamic Gait Index; EDSS: Expanded Disability Status Scale; FSST: Four Square Step Test; MSWS-12: Multiple Sclerosis Walking Scale-12; n: number of patients; PROs: patient-reported outcome measures; PSmob: Performance Scale mobility; pwMS: persons with Multiple Sclerosis; RMI: Rivermead Mobility Index; T25FW: Timed 25 Feet Walking Test; TISmod: Trunk Impairment Scale-modified version; TUG: Timed Up and Go; TUGcogn: Timed Up and Go cognitive; TUGmanual: Timed Up and Go manual

Research

Further research to assess the convergent validity of the mobility measures in pwMS should be done.

Little research has been done concerning the psychometric properties of balance measures in pwMS. This study is the first to analyse the validity of the TUG_{cogn}, FSST, 5STS, mod5STS, TUG_{manual} and TIS-modNV to evaluate mobility in pwMS. Also, not much research has been done about the validity of the BBS and RMI. As such, more research concerning the validity of balance measures in pwMS should be done.

The optional outcome measures should be assessed for more pwMS, leading to more power to draw conclusions about their validity.

The discriminant validity of the optional measures should be evaluated in terms of fallers and non-fallers.

The impact of cognitive disorders should be taken into account when analyzing performance of the PROs.

Further research concerning the reliability, responsiveness and clinical utility of mobility measures (especially balance measures) in pwMS should be done.

Methodological Considerations

There are some remarks concerning the test procedures. There was no randomization of the test order of the measures, although this is important to prevent a carry-over effect. However, this was not done in order to keep testing procedures clear and organized, since a lot of information for each patient had to be documented.

Secondly, to exclude the effect of fatigue, a sufficient time of rest was provided between two tests. However, a better solution would have been that every patient must not perform each measure to prevent fatigue (eg. each patient performs eight of eleven outcome measures).

For part of the ClinROs (T25FW, TUG, TUG_{cogn}, FSST, TUG_{manual}) two trials were performed and the mean of these two trials was used for assessment. In some patients however, only one trial was performed, which could have led to diminished accuracy of performance. On the other hand, performance of two trials could cause a practice effect, as was shown for the T25FW[46].

In the data-analysis, there was no evaluation of the impact of cognition on the outcome measures. Since cognitive disorders may impact the ability to fill out questionnaires and to perform the TUG_{cogn}, this could be an important influencing factor. As such, this should be analyzed in further research.

Conclusion

In general, the overall validity of the ClinROs was higher than the overall validity of the PROs. Correlations between mutual PROs were lower than between mutual ClinROs, which could indicate that the PROs assess different constructs of mobility, and are as such complementary in the information they provide.

Overall, the highest validity was found for the T25FW and 2MWT (ClinROs) and for the PSmob (PRO). For pwMS with a mild disability, the best method to evaluate mobility is to select a combination of ClinROs (2MWT, T25FW) and PROs (PSmob, MSWS-12) to get a more clear picture of the mobility problems. Also for pwMS with a moderate disability, a combination of ClinROs (T25FW, TUG) and PROs (RMI, PSmob) should be used. Although, the ClinROs showed a higher level of validity.
For pwMS with a severe or very severe disability, ClinROs (T25FW, TUG) showed to have the highest validity. Also a combination of outcome measures assessing different aspects of functional mobility (balance, walking, transitions) should be selected.

For the assessment of walking, generally the T25FW and 2MWT are most valid. For the assessment of balance, the TUG_{manual}, TUG and ABC are most valid. And for the assessment of transitions, the TUG_{manual} and TUG are most valid.

As such, an individual selection for each patient should be done, according disability level and the presumably impacted mobility aspect (walking, balance and/or transitions).

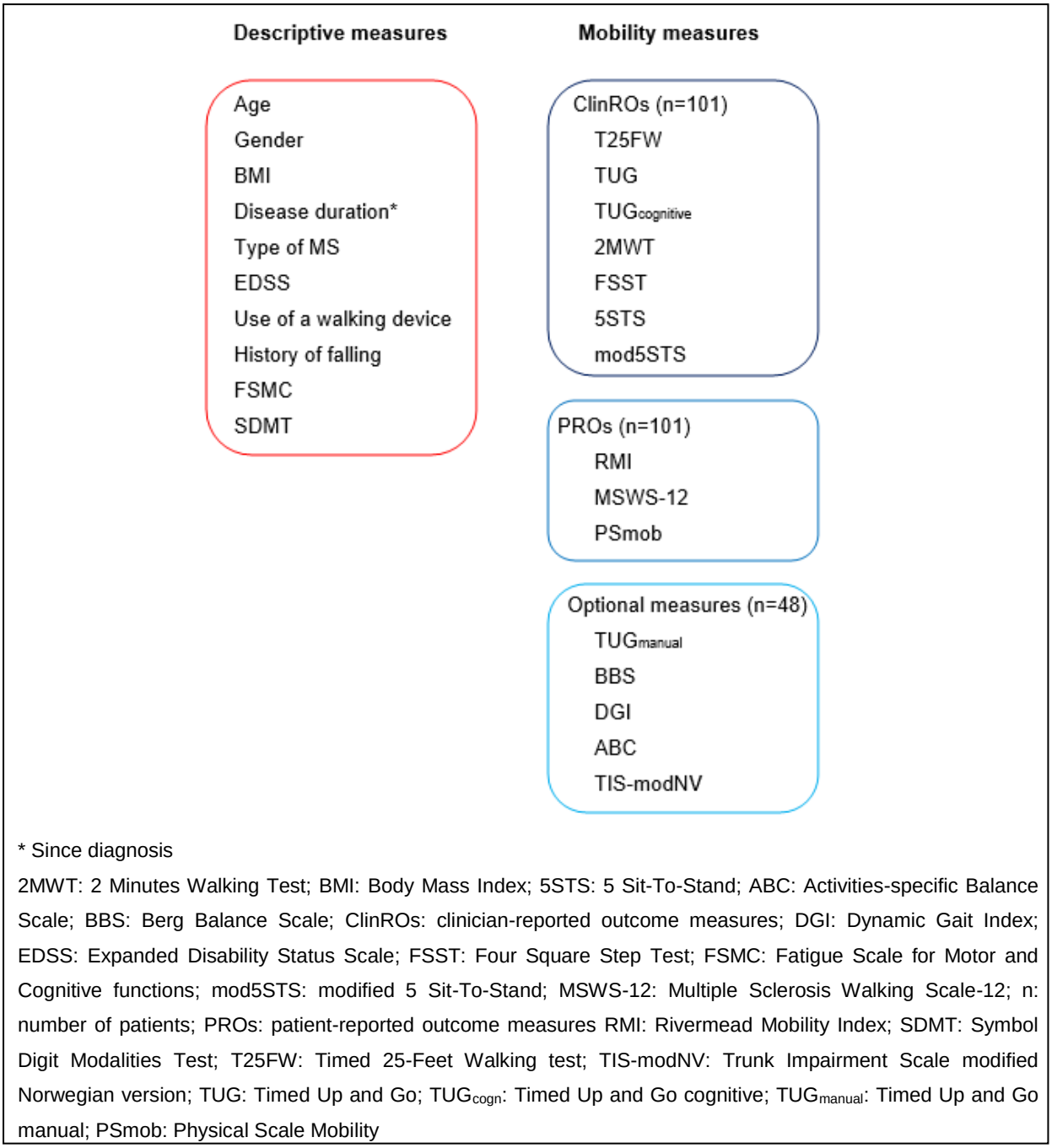
References

1. Larocca NG. Impact of walking impairment in multiple sclerosis: perspectives of patients and care partners. *Patient*. 2011; 4: 189-201.
2. Hobart J, Lamping D, Fitzpatrick R, Riazi A and Thompson A. The Multiple Sclerosis Impact Scale (MSIS-29): a new patient-based outcome measure. *Brain*. 2001; 124: 962-73.
3. Kalron A, Achiron A and Dvir Z. Muscular and Gait Abnormalities in Persons With Early Onset Multiple Sclerosis. *Journal of Neurologic Physical Therapy*. 2011; 35: 164-9.
4. Martin CL, Phillips BA, Kilpatrick TJ, et al. Gait and balance impairment in early multiple sclerosis in the absence of clinical disability. *Multiple sclerosis*. 2006; 12: 620-8.
5. Karst GM, Venema DM, Roehrs TG and Tyler AE. Center of pressure measures during standing tasks in minimally impaired persons with multiple sclerosis. *Journal of neurologic physical therapy : JNPT*. 2005; 29: 170-80.
6. Frzovic D, Morris ME and Vowels L. Clinical tests of standing balance: performance of persons with multiple sclerosis. *ArchPhysMedRehabil*. 2000; 81: 215-21.
7. Cameron MH and Lord S. Postural control in multiple sclerosis: implications for fall prevention. *Current neurology and neuroscience reports*. 2010; 10: 407-12.
8. Cattaneo D, De Nuzzo C, Fascia T, Macalli M, Pisoni I and Cardini R. Risks of falls in subjects with multiple sclerosis. *Archives of physical medicine and rehabilitation*. 2002; 83: 864-7.
9. Peterson EW, Cho CC and Finlayson ML. Fear of falling and associated activity curtailment among middle aged and older adults with multiple sclerosis. *Multiple sclerosis*. 2007; 13: 1168-75.
10. Peterson EW, Cho CC, von Koch L and Finlayson ML. Injurious falls among middle aged and older adults with multiple sclerosis. *Archives of physical medicine and rehabilitation*. 2008; 89: 1031-7.
11. Nilsagard Y, Lundholm C, Denison E and Gunnarsson LG. Predicting accidental falls in people with multiple sclerosis -- a longitudinal study. *ClinRehabil*. 2009; 23: 259-69.
12. Podsiadlo D and Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *JAmGeriatrSoc*. 1991; 39: 142-8.
13. Cattaneo D, Jonsdottir J and Repetti S. Reliability of four scales on balance disorders in persons with multiple sclerosis. *DisabilRehabil*. 2007; 29: 1920-5.
14. Gijbels D, Dalgas U, Romberg A, et al. Which walking capacity tests to use in multiple sclerosis? A multicentre study providing the basis for a core set. *MultScler*. 2012; 18: 364-71.
15. Hobart JC, Riazi A, Lamping DL, Fitzpatrick R and Thompson AJ. Measuring the impact of MS on walking ability: the 12-Item MS Walking Scale (MSWS-12). *Neurology*. 2003; 60: 31-6.
16. Motl RW and Snook EM. Confirmation and extension of the validity of the Multiple Sclerosis Walking Scale-12 (MSWS-12). *JNeurolSci*. 2008; 268: 69-73.
17. Portney LG and Watkins MP. *Foundations of clinical research : applications to practice*. 2nd ed. Upper Saddle River, NJ London: Prentice Hall ; Prentice-Hall International, 2000, p.xiv, 752 p.

18. Kieseier BC and Pozzilli C. Assessing walking disability in multiple sclerosis. *MultScler*. 2012; 18: 914-24.
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. *NeurorehabilNeural Repair*. 2014.
20. Hobart JC, Cano SJ, Warner TT and Thompson AJ. What sample sizes for reliability and validity studies in neurology? *J Neurol*. 2012; 259: 2681-94.
21. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of neurology*. 2011; 69: 292-302.
22. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*. 1983; 33: 1444-52.
23. 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. *Physical therapy*. 2014; 94: 593-608.
24. Fischer JS, Rudick RA, Cutter GR and Reingold SC. The Multiple Sclerosis Functional Composite Measure (MSFC): an integrated approach to MS clinical outcome assessment. National MS Society Clinical Outcomes Assessment Task Force. *Multiple sclerosis*. 1999; 5: 244-50.
25. Butland RJ, Pang J, Gross ER, Woodcock AA and Geddes DM. Two-, six-, and 12-minute walking tests in respiratory disease. *British medical journal*. 1982; 284: 1607-8.
26. Shumway-Cook A, Brauer S and Woollacott M. Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. *Physical therapy*. 2000; 80: 896-903.
27. Dite W and Temple VA. A clinical test of stepping and change of direction to identify multiple falling older adults. *ArchPhysMedRehabil*. 2002; 83: 1566-71.
28. Csuka M and McCarty DJ. Simple method for measurement of lower extremity muscle strength. *The American journal of medicine*. 1985; 78: 77-81.
29. Collen FM, Wade DT, Robb GF and Bradshaw CM. The Rivermead Mobility Index: a further development of the Rivermead Motor Assessment. *IntDisabilStud*. 1991; 13: 50-4.
30. Marrie RA and Goldman M. Validity of performance scales for disability assessment in multiple sclerosis. *Multiple sclerosis*. 2007; 13: 1176-82.
31. Berg KO, Wood-Dauphinee SL, Williams JI and Maki B. Measuring balance in the elderly: validation of an instrument. *CanJPublic Health*. 1992; 83 Suppl 2: S7-11.
32. Shumway-Cook A and Woollacott MH. *Motor control : theory and practical applications*. Baltimore, Md. ; London: Williams & Wilkins, 1995, p.x, 475 p.
33. Powell LE and Myers AM. The Activities-specific Balance Confidence (ABC) Scale. *JGerontolA BiolSciMedSci*. 1995; 50A: M28-M34.
34. Verheyden G, Nieuwboer A, Mertin J, Preger R, Kiekens C and De WW. The Trunk Impairment Scale: a new tool to measure motor impairment of the trunk after stroke. *ClinRehabil*. 2004; 18: 326-34.

35. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. L. Erlbaum Associates, 1988.
36. Portney LG and Watkins MP. *Foundations of clinical research: applications to practice*. Mcgraw-Hill/appleton & Lange, 1993.
37. Ochi H. [Cognitive impairment in multiple sclerosis]. *Brain Nerve*. 2014; 66: 1201-9.
38. Cavanaugh JT, Gappmaier VO, Dibble LE and Gappmaier E. Ambulatory activity in individuals with multiple sclerosis. *JNeurolPhysTher*. 2011; 35: 26-33.
39. Cattaneo D, Regola A and Meotti M. Validity of six balance disorders scales in persons with multiple sclerosis. *DisabilRehabil*. 2006; 28: 789-95.
40. Gijbels D, Eijnde BO and Feys P. Comparison of the 2- and 6-minute walk test in multiple sclerosis. *MultScler*. 2011; 17: 1269-72.
41. Verheyden G, Nuyens G, Nieuwboer A, Van AP, Ketelaer P and De WW. Reliability and validity of trunk assessment for people with multiple sclerosis. *PhysTher*. 2006; 86: 66-76.
42. Moller AB, Bibby BM, Skjerbaek AG, et al. Validity and variability of the 5-repetition sit-to-stand test in patients with multiple sclerosis. *DisabilRehabil*. 2012; 34: 2251-8.
43. Forsberg A, Andreasson M and Nilsagard YE. Validity of the dynamic gait index in people with multiple sclerosis. *PhysTher*. 2013; 93: 1369-76.
44. McGuigan C and Hutchinson M. Confirming the validity and responsiveness of the Multiple Sclerosis Walking Scale-12 (MSWS-12). *Neurology*. 2004; 62: 2103-5.
45. Gijbels D, Alders G, Van HE, et al. Predicting habitual walking performance in multiple sclerosis: relevance of capacity and self-report measures. *MultScler*. 2010; 16: 618-26.
46. Larson RD, Larson DJ, Baumgartner TB and White LJ. Repeatability of the timed 25-foot walk test for individuals with multiple sclerosis. *Clin Rehabil*. 2013; 27: 719-23.

Appendix 1: Summary of study design and outcome measures



* Since diagnosis

2MWT: 2 Minutes Walking Test; BMI: Body Mass Index; 5STS: 5 Sit-To-Stand; ABC: Activities-specific Balance Scale; BBS: Berg Balance Scale; ClinROs: clinician-reported outcome measures; DGI: Dynamic Gait Index; EDSS: Expanded Disability Status Scale; FSST: Four Square Step Test; FSMC: Fatigue Scale for Motor and Cognitive functions; mod5STS: modified 5 Sit-To-Stand; MSWS-12: Multiple Sclerosis Walking Scale-12; n: number of patients; PROs: patient-reported outcome measures RMI: Rivermead Mobility Index; SDMT: Symbol Digit Modalities Test; T25FW: Timed 25-Feet Walking test; 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; PSmob: Physical Scale Mobility

Appendix 2: Overview of the correlations of the outcome measures

Table 1. Correlations of the primary mobility outcome measures (EDSS 2-4).

	T25FW	TUG	TUGcogn	2MWT	FSST	5STS	mod5STS	RMI	MSWS-12	PSmob
ClinROs										
T25FW (n=37)	1.00									
TUG (n=37)	0.78*	1.00								
TUGcogn (n=37)	0.53*	0.79*	1.00							
2MWT(n=37)	-0.87*	-0.84*	-0.68*	1.00						
FSST (n=37)	0.80*	0.75*	0.50*	-0.76*	1.00					
5STS (n=36)	0.58*	0.61*	0.36*	-0.57*	0.67*	1.00				
mod5STS (n=34)	0.73*	0.67*	0.39*	-0.64*	0.71*	0.84*	1.00			
PROs										
RMI (n=37)	-0.57*	-0.65*	-0.58*	0.65*	-0.60*	-0.49*	-0.48*	1.00		
MSWS-12 (n=37)	0.57*	0.56*	0.43*	-0.65*	0.45*	0.38*	0.30	-0.71*	1.00	
PSmob (n=37)	0.52*	0.44*	0.42*	-0.61*	0.43*	0.32	0.28	-0.59*	0.79*	1.00

Note: All non-parametric Spearman's rank correlation coefficients

Abbreviations: *: significant at p<0.05/number of tests; 2MWT: 2 Minutes Walking Test; 5STS: 5 Sit-To-Stand test; ClinROs: clinician-reported outcome measures; FSST: Four Square Step Test; mod5STS: modified 5 Sit-To-Stand test; MSWS-12: Multiple Sclerosis Walking Scale-12; n: number of patients; PROs: patient-reported outcome measures; PSmob: Performance Scale mobility; RMI: Rivermead Mobility Index; T25FW: Timed 25 Feet Walking Test; TUG: Timed Up and Go; TUGcogn: Timed Up and Go cognitive

Table 2. Correlations of the primary mobility outcome measures (EDSS 4.5-5.5).

	T25FW	TUG	TUGcogn	2MWT	FSST	5STS ł	mod5STS ł	RMI ł	MSWS-12 ł	Psmob ł
ClinROs										
T25FW (n=23)	1.00									
TUG (n=23)	0.78*	1.00								
TUGcogn (n=23)	0.49*	0.84*	1.00							
2MWT (n=23)	-0.67*	-0.68*	-0.31	1.00						
FSST (n=21)	0.59*	0.62*	0.34	-0.62*	1.00					
5STS (n=19) ł	0.15	0.15	0.18	0.12	0.04	1.00				
mod5STS (n=18) ł	0.05	0.04	0.02	0.02	-0.09	0.94*	1.00			
PROs										
RMI (n=23) ł	-0.48*	-0.45*	-0.17	0.50*	-0.20	-0.04	-0.16	1.00		
MSWS-12 (n=23) ł	0.21	0.24	-0.08	-0.44*	0.07	0.15	0.32	-0.55*	1.00	
PSmob (n=23) ł	0.62*	0.55*	0.20	-0.49*	0.52*	-0.04	0.04	-0.45*	0.29	1.00

ł: Non-parametric Spearman's rank correlation coefficients

Abbreviations: *: significant at p<0.05/number of tests; 2MWT: 2 Minutes Walking Test; 5STS: 5 Sit-To-Stand test; ClinROs: clinician-reported outcome measures; FSST: Four Square Step Test; mod5STS: modified 5 Sit-To-Stand test; MSWS-12: Multiple Sclerosis Walking Scale-12; n: number of patients; PROs: patient-reported outcome measures; PSmob: Performance Scale mobility; RMI: Rivermead Mobility Index; T25FW: Timed 25 Feet Walking Test; TUG: Timed Up and Go; TUGcogn: Timed Up and Go cognitive

Table 3. Correlations of the primary mobility outcome measures (EDSS 6).

	T25FW	TUG	TUGcogn	2MWT	FSST ł	5STS ł	mod5STS ł	RMI ł	MSWS-12 ł	PSmob ł
ClinROs										
T25FW (n=18)	1.00									
TUG (n=18)	0.91*	1.00								
TUGcogn (n=18)	0.62	0.84*	1.00							
2MWT (n=17)	-0.86*	-0.84*	-0.60	1.00						
FSST (n=16) ł	0.86*	0.79*	0.68	-0.62	1.00					
5STS (n=15) ł	0.40	0.60	0.85*	-0.49	0.27	1.00				
mod5STS (n=16) ł	0.13	0.20	0.34	-0.00	0.01	0.46	1.00			
PROs										
RMI (n=18) ł	-0.48	-0.46	-0.44	0.34	-0.48	-0.24	-0.29	1.00		
MSWS-12 (n=18) ł	0.09	-0.11	-0.08	0.03	0.09	-0.03	0.27	0.11	1.00	
PSmob (n=18) ł	0.34	0.30	0.23	-0.22	0.18	0.37	0.29	-0.54	0.22	1.00

ł: Non-parametric Spearman's rank correlation coefficients

Abbreviations: *: significant at p<0.05/number of tests; 2MWT: 2 Minutes Walking Test; 5STS: 5 Sit-To-Stand; ClinROs: clinician-reported outcome measures; FSST: Four Square Step Test; mod5STS: modified 5 Sit-To-Stand test; MSWS-12: Multiple Sclerosis Walking Scale-12; n: number of patients; PROs: patient-reported outcome measures; PSmob: Performance Scale mobility; RMI: Rivermead Mobility Index; T25FW: Timed 25 Feet Walking Test; TUG: Timed Up and Go; TUGcogn: Timed Up and Go cognitive

Table 4. Correlations of the primary mobility outcome measures (EDSS ≥6.5).

	T25FW	TUG	TUGcogn	2MWT	FSST	5STS	mod5STS	RMI ł	MSWS-12 ł	PSmob ł
ClinROs										
T25FW (n=19)	1.00									
TUG (n=20)	0.91*	1.00								
TUGcogn (n=20)	0.75*	0.87*	1.00							
2MWT (n=20)	-0.93*	-0.85*	-0.72*	1.00						
FSST (n=8)	0.31	0.51	0.50	-0.43	1.00					
5STS (n=10)	0.73*	0.77*	0.60	-0.77*	0.08	1.00				
mod5STS (n=17)	0.84*	0.66*	0.47	-0.75*	0.47	0.61	1.00			
PROs										
RMI (n=21) ł	-0.17	-0.32	-0.27	0.42	-0.12	-0.13	0.04	1.00		
MSWS-12 (n=21) ł	0.42	0.27	0.31	-0.31	0.31	0.36	0.48	0.08	1.00	
PSmob (n=21) ł	0.07	0.17	0.15	-0.13	0.13	0.70*	0.18	-0.21	0.31	1.00

ł: Non-parametric Spearman's rank correlation coefficients

Abbreviations: *: significant at p<0.05/number of tests; 2MWT: 2 Minutes Walking Test; 5STS: 5 Sit-To-Stand; ClinROs: clinician-reported outcome measures; FSST: Four Square Step Test; mod5STS: modified 5 Sit-To-Stand test; MSWS-12: Multiple Sclerosis Walking Scale-12; n: number of patients; PROs: patient-reported outcome measures; PSmob: Performance Scale mobility; RMI: Rivermead Mobility Index; T25FW: Timed 25 Feet Walking Test; TUG: Timed Up and Go; TUGcogn: Timed Up and Go cognitive

Table 5. Correlations of the mobility outcome measures (EDSS 2-4).

	T25FW	TUG	TUGcogn	2MWT	FSST	5STS	mod5STS	TUG manual	BBS	DGI	TIS-mod	RMI	MSWS-12	Psmob	ABC
ClinROs															
T25FW (n=37)	1.00														
TUG (n=37)	0.78*	1.00													
TUGcogn (n=37)	0.53*	0.79*	1.00												
2MWT(n=37)	-0.87*	-0.84*	-0.68*	1.00											
FSST (n=37)	0.80*	0.75*	0.50*	-0.76*	1.00										
5STS (n=36)	0.58*	0.61*	0.36*	-0.57*	0.67*	1.00									
mod5STS (n=34)	0.73*	0.67*	0.39*	-0.64*	0.71*	0.84*	1.00								
TUG manual (n=14)	0.73*	0.95*	0.90*	-0.93*	0.66*	0.35	0.52	1.00							
BBS (n=15)	-0.67*	-0.54*	-0.51	0.78*	-0.52*	-0.18	-0.41	-0.72*	1.00						
DGI (n=14)	-0.64*	-0.65*	-0.60*	0.71*	-0.73*	-0.04	-0.28	-0.74*	0.78*	1.00					
TIS-modNV (n=14)	-0.63*	-0.42	-0.35	0.59*	-0.43	-0.28	-0.17	-0.45	0.44	0.47	1.00				
PROs															
RMI (n=37)	-0.57*	-0.65*	-0.58*	0.65*	-0.60*	-0.49*	-0.48*	-0.60*	0.62*	0.63*	0.28	1.00			
MSWS-12 (n=37)	0.57*	0.56*	0.43*	-0.65*	0.45*	0.38*	0.30	0.66*	-0.50	-0.26	-0.44	-0.71*	1.00		
PSmob (n=37)	0.52*	0.44*	0.42*	-0.61*	0.43*	0.32	0.28	0.82*	-0.79*	-0.65*	-0.55*	-0.59*	0.79*	1.00	
ABC (n=14)	-0.56*	-0.61*	-0.41	0.63*	-0.40	-0.26	-0.59*	-0.62*	0.63*	0.42	0.16	0.51	-0.66*	-0.80*	1.00

Note: All non-parametric Spearman's rank correlation coefficients

Abbreviations: *: significant at p<0.05/number of tests; 2MWT: 2 Minutes Walking Test; 5STS: 5 Sit-To-Stand test; ABC: Activities-specific Balance Scale; BBS: Berg Balance Scale; ClinROs: clinician-reported outcome measures; DGI: Dynamic Gait Index; FSST: Four Square Step Test; mod5STS: modified 5 Sit-To-Stand test; MSWS-12: Multiple Sclerosis Walking Scale-12; n: number of patients; PROs: patient-reported outcome measures; PSmob: Performance Scale mobility; RMI: Rivermead Mobility Index; T25FW: Timed 25 Feet Walking Test; TIS-modNV: Trunk Impairment Scale-modified Norwegian version; TUG: Timed Up and Go; TUGcogn: Timed Up and Go cognitive; TUGmanual: Timed Up and Go manual

Table 6. Correlations of the mobility outcome measures (EDSS >4).

	T25FW	TUG	TUGcogn	2MWT	FSST	5STS	mod5STS	TUG manual	BBS	DGI	TIS-mod	RMI	MSWS-12	Psmob	ABC
ClinROs															
T25FW (n=60)	1.00														
TUG (n=61)	0.91*	1.00													
TUGcogn (n=61)	0.76*	0.90*	1.00												
2MWT (n=60)	-0.87*	-0.84*	-0.68*	1.00											
FSST (n=45)	0.69*	0.70*	0.61*	-0.70*	1.00										
5STS (n=44)	0.45*	0.57*	0.60*	-0.38*	0.27	1.00									
mod5STS (n=51)	0.57*	0.60*	0.57*	-0.49*	0.30	0.82*	1.00								
TUG manual (n=24)	0.86*	0.94*	0.69*	-0.83*	0.71*	0.50*	0.46*	1.00							
BBS (n=33)	-0.82*	-0.76*	-0.67*	0.70*	-0.45*	-0.65*	-0.78*	-0.63*	1.00						
DGI (n=32)	-0.68*	-0.65*	-0.65*	0.60*	-0.44*	-0.54*	-0.66*	-0.50*	0.82*	1.00					
TIS-modNV (n=33)	-0.57*	-0.49*	-0.46*	0.49*	-0.27	-0.49*	-0.47*	-0.28	0.77*	0.73*	1.00				
PROs															
RMI (n=62)	-0.51*	-0.52*	-0.43*	0.52*	-0.47*	-0.21	-0.34*	-0.26	0.20	0.11	0.13	1.00			
MSWS-12 (n=62)	0.34*	0.25	0.16	-0.32*	0.14	0.19	0.36*	0.40	-0.45*	-0.46*	-0.25	-0.18	1.00		
PSmob (n=62)	0.61*	0.65*	0.54*	-0.56*	0.49*	0.39*	0.45*	0.44*	-0.62*	-0.46*	-0.35*	-0.52*	0.33*	1.00	
ABC (n=33)	-0.64*	-0.54*	-0.49*	0.55*	-0.30	-0.56*	-0.71*	-0.29	0.63*	0.59*	0.41*	0.27	-0.45*	-0.30	1.00

Note: All non-parametric Spearman's rank correlation coefficients

Abbreviations: *: significant at p<0.05/number of tests; 2MWT: 2 Minutes Walking Test; 5STS: 5 Sit-To-Stand test; ABC: Activities-specific Balance Scale; BBS: Berg Balance Scale; ClinROs: clinician-reported outcome measures; DGI: Dynamic Gait Index; FSST: Four Square Step Test; mod5STS: modified 5 Sit-To-Stand test; MSWS-12: Multiple Sclerosis Walking Scale-12; n: number of patients; PROs: patient-reported outcome measures; PSmob: Performance Scale mobility; RMI: Rivermead Mobility Index; T25FW: Timed 25 Feet Walking Test; TIS-modNV: Trunk Impairment Scale-modified Norwegian version; TUG: Timed Up and Go; TUGcogn: Timed Up and Go cognitive; TUGmanual: Timed Up and Go manual

Appendix 3: Overview of convergent validity of the outcome measures

Table 7. Overview of convergent validity of the primary outcome measures

Outcome measure	Total sample (n=101)	EDSS 2-4 (n=37)	EDSS 4.5-5.5 (n=23)	EDSS 6 (n=18)	EDSS ≥ 6.5 (n=21)
ClinRO					
T25FW	moderate to excellent	moderate to good	poor to good	poor to excellent	poor to excellent
TUG	fair to excellent	fair to good	poor to good	poor to excellent	poor to excellent
TUGcogn	fair to excellent	fair to good	poor to good	poor to good	poor to good
2MWT	moderate to excellent	moderate to excellent	poor to moderate	poor to good	poor to excellent
FSST	fair to good	fair to good	poor to moderate	poor to good	poor to moderate
5STS	fair to good	poor to good	poor (to excellent)	poor to good	poor to good
mod5STS	fair to good	poor to good	poor (to excellent)	poor to fair	poor to good
PRO					
RMI	fair to good	fair to moderate	poor to moderate	poor to moderate	poor to fair
MSWS-12	fair to moderate	fair to good	poor to moderate	poor to fair	poor to fair
PSmob	moderate to good	fair to good	poor to moderate	poor to moderate	poor to moderate

Abbreviations: 2MWT: 2 Minutes Walking Test; 5STS: 5 Sit-To-Stand; EDSS: Expanded Disability Status Scale; FSST: Four Square Step Test; mod5STS: modified 5 Sit-To-Stand test; MSWS-12: Multiple Sclerosis Walking Scale-12; n: number of patients PSmob: Performance Scale mobility; RMI: Rivermead Mobility Index; T25FW: Timed 25 Feet Walking Test; TUG: Timed Up and Go; TUGcogn: Timed Up and Go cognitive

Table 8. Overview of convergent validity of the secondary outcome measures

Outcome measure	Total sample (n=48)	EDSS ≤4 (n=15)	EDSS >4 (n=33)
ClinRO			
TUGmanual	fair to excellent	fair to excellent	fair to excellent
BBS	fair to good	poor to good	poor to good
DGI	fair to good	poor to good	poor to good
TIS-modNV	fair to good	poor to moderate	fair to good
PRO			
ABC	fair to moderate	fair to good	fair to moderate

Abbreviations: ABC: Activities-specific Balance Scale; BBS: Berg Balance Scale; DGI: Dynamic Gait Index; EDSS: Expanded Disability Status Scale; n: number of patients; TIS-modNV: Trunk Impairment Scale-modified Norwegian version; TUGmanual: Timed Up and Go manual

Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
Validity of mobility measures in multiple sclerosis, according disability level: a RIMS multi-center study.

Richting: **master in de revalidatiewetenschappen en de kinesitherapie-revalidatiewetenschappen en kinesitherapie bij musculoskeletale aandoeningen**
Jaar: **2015**

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Voor akkoord,

Jans, Laura **Vanoppen, Kristien**