Psychometric properties of the modified reaching performance scale in persons with multiple sclerosis

Peer-reviewed author version

RAATS, Joke; FEYS, Peter; GYSEMBERG, Griet; Ferdinand, S; Levin, MF & LAMERS, Ilse (2024) Psychometric properties of the modified reaching performance scale in persons with multiple sclerosis. In: Multiple Sclerosis and Related Disorders, 90 (Art N° 105806).

DOI: 10.1016/j.msard.2024.105806
Handle: http://hdl.handle.net/1942/46283

Psychometric properties of the modified Reaching Performance Scale in persons with Multiple Sclerosis.

Raats J[1,2], Feys P [1,2], Gysemberg G[1,3], Ferdinand S[4], Levin MF[5], Lamers I [1,2,3]

1 UMSC, University MS Centre, Hasselt Pelt, Belgium
2 REVAL Rehabilitation Research Centre, Faculty of Rehabilitation Science, UHasselt, Belgium
3 Noorderhart, Rehabilitation and MS Center, Pelt, Belgium
4 National MS Center, Melsbroek, Belgium
5 School of Physical and Occupational Therapy, Faculty of Medicine and Health Sciences, McGill University, Montreal, Canada

Abstract

Background
A valid and reliable assessment tool to describe the quality of the movement pattern of reaching can provide valuable insights into motor performance deficits in persons with MS (pwMS). The Reaching Performance Scale, developed for stroke, is a promising scale to assess movement patterns in pwMS. However, psychometric properties of the scale are lacking in pwMS.

Objectives
Firstly, to investigate the content validity of the modified Reaching Performance Scale for application in patients with MS (mRPS). Secondly, to investigate the psychometric properties (within- and between-session reliability and concurrent validity) of the mRPS for pwMS.

Methods
Forty-five pwMS (mean EDSS 6.6 pt, IQR 6-7.5) executed the mRPS that rates the quality of movement patterns and compensations during reach to grasp tasks. The content validity was determined by an expert panel based on observations of subjects performing the RPS. The reliability was based on five repetitions within one day, and between two days. For the concurrent validity, outcome measures at two levels of the International Classification of Functioning were correlated with the mRPS: Body Structure and Function level: Fugl-Meyer Assessment of the Upper Limb (FMA-UL), maximal isometric hand grip strength (HGS; Activity level: Action Research Arm Test (ARAT), Box and Blocks Test (BBT), Nine Hole Peg Test (NHPT) and Trunk Impairment Scale 2.0 (TIS 2.0) as well as perceived performance by the Manual Ability Measure-36 (MAM-36) .

Results
Scale modifications were made only on the ratings of the trunk displacement subscale. The mRPS had excellent agreement scores for within-session reliability (range of Kappa between 0.85 and 0.98) and moderate-to-excellent agreement scores for between-session reliability (K: 0.66-1.00). Regarding validity, the mRPS was highly correlated with the ARAT (rho=0.74, p<0.001), followed by moderate correlations with trunk performance (TIS 2.0, rho= 0.61, p<0.001), hand function (BBT: rho=0.64, p<0.001; NHPT: rho=-0.61, p<0.001) and perceived performance (MAM36 rho= 0.53, p<0.001).

Conclusion
The mRPS is a reliable measurement tool to describe the movement pattern quality and motor compensations used during reaching in pwMS. Concerning concurrent validity, the mRPS is partially related to other measures of upper limb and trunk performance.

Introduction

Reaching and grasping are integral to the performance of activities of daily life. Daily activities consist of task-dependent movements with the goal of bringing the endpoint (hand) to a desired location within or beyond the arm workspace and at a desired time, considering physical, task and environmental constraints [1-5]. In people with neurological lesions, movement patterns used for reaching and grasping are disrupted. The capacity to produce many different movement patterns to accomplish the same action goal is reduced (i.e., motor equivalence). Compensations are new motor patterns that arise from substitution for lost motor elements or the adaptation of remaining motor elements, such that functions are taken over or modified by the incorporation of different end effectors or body segments [6], i.e. compensatory patterns of shoulder and trunk movements during reaching [7-11]. For example, common motor compensations used for reaching are excessive forward trunk movement in patients with stroke [12] and backward trunk movement, as a counterbalance, in patients with spinal cord injury [11]. There is little information about reaching deficits in persons with Multiple Sclerosis (pwMS). Sixty percent of pwMS reported having problems with their upper limb movements [13], in particular with reaching and grasping. Upper limb dysfunction can be mapped by a broad spectrum of outcome measures as the nine hole peg test (NHPT), box and block test (BBT) or action research arm test (ARAT). Most measures used in MS primarily quantify the degree of task completion on ordinal scales and/or by time, without considering movement quality. Movement quality variables include measures of range of joint movement, spatial and temporal interjoint coordination, and muscle activation patterns [7, 14]. Analyzing movement patterns used for reaching in pwMS could provide valuable information about compensatory or maladaptive movements [3, 7] and can be used to personalize therapy content.

Movement kinematics are the gold standard to assess movement quality and are increasingly recommended in research [15]. However kinematic assessments are not yet commonly used in the clinical field because of the historical high cost of the equipment and the complexity of their analysis and interpretation. In recent years, low cost tracking methodologies are being developed for applications in neurological conditions but products are not available 'off the shelf' yet, especially not for measuring trunk and upper limb movements. An alternative approach is to use detailed clinical assessment based on observational kinematics, wherein the clinician visually identifies and quantifies different features of the movement [16]. One outcome measure that relies on observational kinematics to specifically describe different motor components including motor compensations during reach-to-grasp tasks is the Reaching Performance Scale for Stroke (RPSS)[17]. Six different movement components are analyzed during reaching performance to targets placed close to and far from the body. The first four movement components of the test score trunk displacement, endpoint smoothness, and shoulder and elbow movements. The last two movement components evaluate the grasp function and overall quality of the task performance. Four-point ordinal scales are used to rate each of the different components. The total score on the RPSS ranges from 0 to 18 both for reaching to a close and far target, rated separately. A higher score indicates a better performance. The RPSS has acceptable intra-rater (Close target: ICC 0.80-0.95 (0.50-0.98); Far target: ICC 0.83- 0.94 (0.56-0.98) and inter-rater reliability (Close target: ICC 0.84 (0.58-0.95); Far target: ICC 0.89 (0.69-0.96) in people with stroke [17]. Specific for post-stroke patients with a chronic UL hemiparesis, excellent test-retest reliability was shown for close (ICC = 0.98, 95% confidence interval [CI] 0.94–0.99) and far targets (ICC = 0.98, 95% CI 0.95–0.99) [18].Based on the heterogenic patterns in reaching in different neurological pathologies [11, 12], it is plausible that pwMS may show different reaching patterns and/or impairments compared to people with stroke. Stroke-related trunk control deficits often reflect localized damage to specific brain areas, resulting in unilateral impairments and asymmetrical posture such as leaning towards the hemiparetic side. MS-related trunk control deficits may stem from more distributed or diffuse central nervous system damage, leading to bilateral and more symmetrical postural changes as increased sway and difficulty maintaining an upright posture. The different trunk control deficits in pathologies are thought to impact differently on trunk stability, alignment and consequently production of upper limb reaching movements. All may lead to compensatory movements. For adequate rehabilitation in MS,it is crucial to have valid, reliable and easy to use outcome measures to capture movement patterns and compensations during reaching.

Therefore this study aimed 1) to modify the content of the RPS for application in pwMS; and 2) to investigate the psychometric properties (within- and between-session reliability and concurrent validity) of the modified Reaching Performance Scale (mRPS) for pwMS.

3. Methods

3.1. Procedure

This cross-sectional observational design in pwMS was approved by the Medical Ethics Committee of the Ziekenhuis Oost-Limburg (number: B371201838357/NCT04908891) and the respective ethics committees of the participating hospitals (Nooderhart Rehabilitation & MS Center (R&MSC) Pelt and National MS Center Melsbroek).

Recruitment and a first screening of patients was done in collaboration with therapists. Participants were informed about the protocol and gave written consent. All participants were older than 18, able to understand and execute the test instructions, and able to move at least one out of three upper limb joints (wrist, elbow or shoulder) against gravity. Participants were excluded if they had other medical conditions interfering with upper limb or trunk function, such as orthopedic or other neurologic conditions/ pathology, perceived pain in one of the upper limbs of >4/10 measured with the Visual Analog Scale, severe cognitive or visual deficits interfering with testing or a relapse or relapse-related treatment in the last month prior to the study.
The study consisted of 2 phases. In the first phase, the modified RPS was formulated by an expert group of 6 individuals and in the second phase, the psychometric properties of the mRPS were determined.

3.2. Descriptive measures

The following data were collected to describe the study population: sex, age, type of MS, disease duration and Expanded Disability Status Scale (EDSS) [19] obtained from the medical record. Hand dominance was determined with the Edinburgh Handedness Inventory [20]. Active and passive range of motion of shoulder flexion, elbow extension and wrist flexion and extension was measured with goniometry. The Hauser Ambulation Index (HAI) was used to assess mobility and the extent of support necessary to walk. A higher score (0-10 pt) indicates a higher need of support. In addition, the Modified Fatigue Impact Scale (score between 0-84 pt) was used to obtain a general overview of the participant's perceived fatigue[21]. Cognitive function was evaluated with Symbol Digit Modalities Test [22].

3.3. Modified Reaching Performance Scale: data collection

The modified Reaching Performance Scale was used to evaluate reaching performance to close and far targets from a sitting position. The task involved reaching and grasping a cardboard cone placed on the table in the midline of the trunk, 1 cm (close target) and 30 cm (far target) from the front edge of the table. The maximal total score on the mRPS is 36, 18 points for reaching to a close target ad 18 points to reaching to a far target. The reaching performance was rated based on video viewing of the movement. One camera was placed in front of the participant and another was placed lateral from the tested side. The full test set up and the modified Reaching Performance Scale can be found in Appendix 1.

3.4. Upper limb and trunk outcome measures

Based on the different levels of the International Classification of Functioning, different outcome measures were selected according to Body Structure and Function: Fugl-Meyer Assessment of the Upper Limb (FMA-UL), Maximal isometric hand grip strength (HGS), and to Activity level: Action Research Arm Test (ARAT), Box and Blocks Test (BBT), Nine Hole Peg test (NHPT), Manual Ability Measure-36 (MAM-36) and the Trunk Impairment Scale 2.0 (TIS 2.0). All upper limb assessments were performed bilaterally and the order of testing was randomized between participants.
The FMA-UL was used to assess upper limb motor function [23]. The upper limb section consists of 33 items rated on 3-point ordinal scales for a total score of 66 points, which indicates normal motor function. The reliability of the FMA-UL is excellent. The HGS was established by using hand held dynamometry [24]. The average hand grip strength (kg) was measured with the JAMAR® hand-held dynamometer during 3 repeated trials, using the procedure recommended by the American Society of Hand Therapists [25]. The HGS has excellent interrater reliability [26]. The NHPT was used to evaluate manual dexterity by recording the time needed to place and then remove nine pegs in a pegboard was recorded and averaged over two trials [27]. The ARAT was used to evaluate unilateral upper limb function. The test consists of 19 items, with four subtests (grasp, grip, pinch and gross movement), which are scored on 4-point ordinal scales with a maximum score of 57 points. A higher score indicates better performance. The reliability of the ARAT is excellent [23]. The BBT assesses gross manual dexterity by having the patient lift and transport blocks from one side of a box to another. The maximal number of blocks transported in one minute is recorded. The BBT demonstrates excellent test-retest and interrater reliability and is strongly correlated with the ARAT and the FMA-UL [23]. The TIS 2.0 was used to assess dynamic sitting balance and trunk coordination in sitting. Each item is scored using a 2- or 3-point ordinal scale and the total score ranges from 0 to 16 points, where a higher score indicates better performance. Interrater and test-retest agreement was excellent [28].
Perceived manual ability while performing unimanual and bimanual tasks was evaluated with the MAM-36. Thirty-six items were rated using 4-point ordinal scales, with a low score indicating high perceived disability [29-31].

3.5. Phase 1: Content validity

The methodology described by Hasson et al. [32] was used to determine the content validity of the scale. First, the research problem and rationale were defined as 'What are clinical instruments to quantify the quality of a reaching performance'. Secondly, a search of the literature was performed to identify clinical outcome measures of reaching performance that specifically assess movement quality. The Reaching Performance Scale for Stroke was chosen. Our findings were in line with a systematic review conducted by

Demers and Levin [7]. They identified outcome measures used in neurological patients to assess UL movement quality. In the next phase, the findings were shared with an expert panel specifically selected based on either their extensive clinical knowledge of the population, either research expertise on upper limb/trunk assessment in neurological populations or the development of the original scale. The expert group was composed of 6 individuals (1 occupational therapist and 5 physical therapists). All participants were academic researchers, with extensive clinical previous experience or clinical experts with research understanding. They all lead research or clinical programs in trunk and/or upper limb mobility in neurological patients (stroke and MS). They were informed of the research problem by a presentation. Then they were asked to rate the video recordings of the reaching movements made by the pwMS using the RPS. During the first session, they were asked to apply the original scale and rating descriptions of the RPS and provide remarks on different items. Based on an iterative process of 4 sessions with video viewing and discussion, a group consensus was arrived at and the final modifications of the scale description were determined. An addition to the scale was formulated on the subscale of trunk displacement. Specifically, a second movement pattern description was added with signs of mild or marked signs of imbalance and no trunk displacement during reaching.

3.6. Phase 2. Within- and between-session reliability and concurrent validity

Two test sessions in which pwMS were asked to perform the mRPS reaching tasks were scheduled within one week with a maximal duration of one hour each. The sessions occurred at the same time of day, with a maximal variance of two hours to avoid the impact of diurnal fatigue. All descriptive and secondary upper limb and trunk outcome measures were collected in a random order during the two test sessions. For the execution of the mRPS, 5 repetitions of the reaching movement to each of the close and far targets were executed on both days. For the within-session reliability, only the data collected in the first session were analyzed. For the between-session reliability, the second execution of the mRPS of each day was used. For the concurrent validity of the mRPS, different upper limb and trunk outcome measures were executed divided between the two different sessions, tested in a random order: FMA-UL, ARAT, BBT, NHPT, MAM36, TIS 2.0.

3.7. Statistical analysis

Data were analysed using SPSS software version 25. For all data, normality was checked with the Kolmogorov-Smirnov tests and homogeneity of variances were verified by Levene's tests. To examine relative reliability, agreement scores were calculated by kappa coefficients and percentage agreement. Kappa values <0.40 indicated 'poor', between 0.41 and 0.60 indicated 'moderate', and between 0.61 and 0.80 indicated 'good' agreement. Values > 0.81 were considered as 'excellent' [33]. To examine absolute reliability, the SEM (standard error of measurement) and MDC (minimal detectable change) were calculated for reaching to a close and far target. Significance levels were set at $p < 0.05$. It was aimed to recruit 50 participants as recommended by the COSMIN guidelines.[34]

To investigate the concurrent validity of the mRPS with the upper limb and trunk outcome measures (FMA-UL, NHPT, ARAT, BBT, MAM36, TIS 2.0) Pearson or Spearman correlation coefficients were calculated, as indicated. Correlations were interpreted as poor ($r \leq 0.30$), weak ($r=0.30$-0.50), moderate ($r= 0.50$-0.70), strong ($r=0.70$-0.90) or very strong ($r \geq 0.90$). We present the data for the right hand with results being similar for the left hand.

4. Results

4.1. Patient characteristics

A total of 50 pwMS were screened according to inclusion/exclusion criteria. Four persons were excluded due to a recent shoulder infiltration, shoulder fracture, a relapse or severe cognitive problems. Of the remaining 46 patients included in the study, 45 pwMS participated in both testing sessions. Subsequently, data from 45 pwMS were analyzed (Table 1).

4.2. Content validity

Based on the videos, the expert group agreed that overall, the proposed items and wording of the original scale should be retained. For the subscale of trunk displacement, additional formulations were needed, since pwMS could display movement patterns that were not captured by the descriptions used in the original version of the scale. Thus, on this subscale, different adaptations of trunk displacement, specific to pwMS were added. For example, during reaching towards the far target, a small trunk displacement is observed in healthy subjects. However, two different patterns of trunk displacement were identified in our neurological study sample and therefore the descriptions of the trunk displacement were modified. The first pattern was an excessive, compensatory, forward displacement of the trunk. The second pattern was a limited forward trunk displacement or an unstable trunk displacement. Examples of different compensation strategies of an unstable trunk displacement were grasping to the wheelchair armrest for support, using the other arm to counterbalance, hyperextending the neck to counterbalance. The modified Reaching Performance Scale (mRPS) is presented in Appendix 1.

4.3. Within-session reliability of mRPS
Table 2 shows the within-session reliability of the mRPS. For one person who was only able to execute the mRPS once, trials 2/3/4 and 5 are missing.
Excellent agreement was found for reaching to the close and the far targets between all different trials (k = 0.83 to 0.98) with percentage agreement scores above 90% for both the close and far targets.

4.4. Between-session reliability of the mRPS
Table 3 shows the between-session reliability of the mRPS. Moderate to excellent kappa values were found for reaching to the close target (range between 0.57-1.00) and the far target (range between 0.59-0.80). The percentage agreement scores were excellent: 82.20 and 100%. The SEM (MDC) for reaching to close and far target were, respectively, 2.27 (6.29) and 1.36 (3.78).

4.5. Concurrent validity of the mRPS
Table 4 shows the concurrent validity of the mRPS with the secondary upper limb outcome measures. Significant correlations were found between the mRPS- close target and mRPS-far target and the TIS 2.0, AROM of the shoulder, hand grip strength, ARAT, BBT, NHPT and the MAM36. The lowest correlation was found between TIS 2.0 and mRPS- close target (rho= 0.38, p<0.001). Higher correlations were detected with gross motor tasks (ARAT and BBT) compared to a dexterity task (NHPT). Furthermore, there was a moderate correlation between the mRPS- far target and perceived upper limb performance (MAM36).

5. Discussion

This study investigated the content validity of the Reaching Performance Scale (RPS) for application in patients with MS (mRPS), leading to a modified Reaching Performance Scale (mRPS). Secondly psychometric properties (within- and between-session reliability and concurrent validity) of this modified mRPS for pwMS were investigated.

5.1. Content validity
One item on the original RPS, developed for patients with stroke, needed to be modified for the MS population because it did not adequately capture the variability in trunk performance in this patient group during reaching. Two main patterns were observed by the experts: one to compensate for deficits in UL movements by an excessive forward displacement, and one to compensate for the deficits in trunk control. For example, we observed that persons were anxious about falling against the table during reaching to the far target. Patients created a counterbalancing force by swinging out the other arm, by supporting their trunk by grasping the wheelchair armrest with the other hand or by executing the reaching movement very fast due to impaired trunk control. These findings confirm that trunk control is crucial to perform reaching movement. In healthy individuals, compensatory and anticipatory postural adjustments are needed to maintain stability when posture is disturbed by movements of the upper limb [11]. The diverse patterns shown by pwMS were in contrast to the more stereotypical compensatory patterns shown by people with stroke (i.e. excessive trunk displacement during forward reaching). This could be explained by the greater heterogeneity of lesion locations in pwMS and higher prevalence of bilateral trunk and upper limb impairments [35]. Postural control is achieved by an interaction of different systems and brain areas in the central nervous system. The reticulo- and vestibulospinal tracts play crucial roles in the activation of postural muscles, the medial corticospinal tract in the activation of the trunk muscles and the brainstem tracts in anticipatory and compensatory postural adjustments [11]. It is likely that the pwMS have lesions in more than one of these regions, which is less likely in persons with stroke.

5.2. Within- and between-session reliability
The excellent and good agreement scores for the between- and within-session reliability indicate that the mRPS is a reliable measurement tool to evaluate reaching movements in pwMS. The excellent within-session agreement values indicate that executing 5 reaching trials to each target during the mRPS does not cause excessive fatigue of the upper limb, nor learning. This suggests that the included patient with MS were stable in executing the reaching movement within one session. These results are in line with the reliability documented in stroke [18]. One must however note the limitation that our results in MS were ultimately obtained from 45 participants while minimally 50 is recommended by the COSMIN guidelines.[34]

5.3. Concurrent validity
It was hypothesized that sensorimotor impairments in the trunk would have an impact on reaching performance. This hypothesis is supported by the moderate to good correlations between the different

scores on the mRPS (close target, far target and total score) and the TIS 2.0. The higher correlation for the mRPS-far target indicates a more active contribution of the trunk compared with the mRPS-Close Target.
The moderate correlations between mRPS and upper limb outcome measures confirms the hypothesis, specifically that the mRPS captures motor performance of the upper limb, not evaluated by other measures. The good correlation between the mRPS–Far Target and the ARAT can be explained by the fact that the ratings during the upper limb performance of the ARAT also includes, to a limited extent, the quality of movement during reaching and grasping tasks. A score of 2 is given in the ARAT when the patient shows compensatory movements. The findings obtained in pwMS are in line with the correlations found in Levin et al.[17] in stroke patients. The original scale of the RPS has good to excellent correlations with the TEMPA (range between r= 0.84-0.88). Nevertheless, a poor correlation was seen between the mRPS and the grip force (rho= 0.35, p<0.001), which is in contrast to the findings reported for the original RPS, where a moderate correlation was reported (r=0.60 and r=0.63) [17]. The difference could be explained by the different pathophysiology of the diseases, including the heterogeneity with potentially lesions in multiple brain regions and specific tracts in pwMS compared to people with stroke.

Furthermore, we found a moderate correlation (r= 0.55) between the mRPS-far Target and perceived upper limb performance (MAM36), which indicates that the mRPS is related to the ability to execute tasks with the upper limb during daily life.

5.4. Relevance for the clinical field
From a clinimetric perspective, the mRPS can be used as an informative clinical tool to assess reaching movement patterns, and observe potential compensatory movements in pwMS. The mRPS is easy and fast to execute and reaching performance can be rated in a simple way. To the best of our knowledge, the mRPS is the only outcome measure that includes an assessment of trunk movement during an UL movement. Specifically, reaching to a far target requires an active recruitment of the trunk, and consequently also provides an indication of the trunk performance of pwMS, which is known to be often impaired [36]. Assessing the whole chain (trunk, shoulder, elbow, grasping by the hand) in a comprehensive way during reaching in pwMS can provide valuable information, not captured by traditional clinical scales. This information can be used to optimize and personalize rehabilitation programs targeting trunk and upper limb function [37]. Regarding psychometric research, future research on the mRPS should investigate responsiveness of the mRPS in MS, in either a longitudinal or rehabilitation perspective. For now, data on measurement error were provided by means of standard error of measurement and the minimal detectable change.
Regarding generalizability, we conclude that a range of upper limb impairment severity was present in our sample to reflect the MS population [35]. However, it is noted that the sample had on average so-called severe disability as reflected by the interquartile range of the EDSS being 6 to 7.5, or meaning needing assistive devices to walk 20 meters. Subjects were not recruited based on EDSS, but the recruitment via rehabilitation centers may have contributed to this bias. Another element is that upper limb as well as trunk impairment is more prevalent or severe in pwMS in the higher EDSS range [35, 36]. In line with the subject of upper limb impairment, and recruitment routes, we note that the majority of the sample was older than 50 years.
A limitation of the study was the sample size that did not allow to fully take heterogeneity of the severity of upper limb and trunk dysfunction or underlying impairments such as hypertonia or muscle weakness into account. In addition, one may comment that the mRPS could be influenced by symptoms fatigability. Given that the execution of the mRPS can be considered as a low intensity exercise, it should not impact maximal motor capacity [38]. We can however not exclude the possibility that the repeated required performances related to the number of different items, could have led to perceived fatigue.

6. Conclusion
Using the mRPS, it is possible to reliably rate the trunk and upper limb movement patterns in pwMS while reaching to a close and far target. Based on the concurrent validity analyses, it is indicated that the trunk, shoulder and upper limb all play a crucial roles in reaching performance.

7. References

1. Rosenbaum, D.A., et al., *The problem of serial order in behavior: Lashley's legacy.* Hum Mov Sci, 2007. **26**(4): p. 525-54.
2. Latash, M.L., et al., *Structure of motor variability in marginally redundant multifinger force production tasks.* Exp Brain Res, 2001. **141**(2): p. 153-65.

3. Tomita, Y., M.R.M. Rodrigues, and M.F. Levin, *Upper Limb Coordination in Individuals With Stroke: Poorly Defined and Poorly Quantified.* Neurorehabil Neural Repair, 2017. **31**(10-11): p. 885-897.
4. Mandon, L., et al., *Faster Reaching in Chronic Spastic Stroke Patients Comes at the Expense of Arm-Trunk Coordination.* Neurorehabil Neural Repair, 2016. **30**(3): p. 209-20.
5. Shaikh, T., et al., *Arm-trunk coordination for beyond-the-reach movements in adults with stroke.* Neurorehabil Neural Repair, 2014. **28**(4): p. 355-66.
6. Levin, M.F., J.A. Kleim, and S.L. Wolf, *What do motor "recovery" and "compensation" mean in patients following stroke?* Neurorehabil Neural Repair, 2009. **23**(4): p. 313-9.
7. Demers, M. and M.F. Levin, *Do Activity Level Outcome Measures Commonly Used in Neurological Practice Assess Upper-Limb Movement Quality?* Neurorehabil Neural Repair, 2017. **31**(7): p. 623-637.
8. Tomita, Y., A.A. Mullick, and M.F. Levin, *Reduced Kinematic Redundancy and Motor Equivalence During Whole-Body Reaching in Individuals With Chronic Stroke.* Neurorehabil Neural Repair, 2018. **32**(2): p. 175-186.
9. Cirstea, M.C. and M.F. Levin, *Compensatory strategies for reaching in stroke.* Brain, 2000. **123 (Pt 5)**: p. 940-53.
10. Levin, M.F., et al., *Compensatory Versus Noncompensatory Shoulder Movements Used for Reaching in Stroke.* Neurorehabil Neural Repair, 2016. **30**(7): p. 635-46.
11. Peeters, L.H.C., I.J.M. de Groot, and A.C.H. Geurts, *Trunk involvement in performing upper extremity activities while seated in neurological patients with a flaccid trunk - A review.* Gait Posture, 2018. **62**: p. 46-55.
12. Levin, M.F., et al., *Use of the trunk for reaching targets placed within and beyond the reach in adult hemiparesis.* Exp Brain Res, 2002. **143**(2): p. 171-80.
13. Kister, I., et al., *Natural history of multiple sclerosis symptoms.* Int J MS Care, 2013. **15**(3): p. 146-58.
14. Lamers, I., et al., *Upper limb assessment in multiple sclerosis: a systematic review of outcome measures and their psychometric properties.* Arch Phys Med Rehabil, 2014. **95**(6): p. 1184-200.
15. Kwakkel, G., et al., *Standardized Measurement of Quality of Upper Limb Movement After Stroke: Consensus-Based Core Recommendations From the Second Stroke Recovery and Rehabilitation Roundtable.* Neurorehabil Neural Repair, 2019. **33**(11): p. 951-958.
16. Bernhardt, J., P.J. Bate, and T.A. Matyas, *Accuracy of observational kinematic assessment of upper-limb movements.* Phys Ther, 1998. **78**(3): p. 259-70.
17. Levin, M.F., et al., *Development and validation of a scale for rating motor compensations used for reaching in patients with hemiparesis: the reaching performance scale.* Phys Ther, 2004. **84**(1): p. 8-22.
18. Subramanian, S.K., et al., *Reaching performance scale for stroke - Test-retest reliability, measurement error, concurrent and discriminant validity.* PM R, 2022. **14**(3): p. 337-347.
19. Kurtzke, J.F., *Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS).* Neurology, 1983. **33**(11): p. 1444-52.
20. Ransil, B.J. and S.C. Schachter, *Test-retest reliability of the Edinburgh Handedness Inventory and Global Handedness preference measurements, and their correlation.* Percept Mot Skills, 1994. **79**(3 Pt 1): p. 1355-72.
21. Learmonth, Y.C., et al., *Psychometric properties of the Fatigue Severity Scale and the Modified Fatigue Impact Scale.* J Neurol Sci, 2013. **331**(1-2): p. 102-7.
22. Van Schependom, J., et al., *The Symbol Digit Modalities Test as sentinel test for cognitive impairment in multiple sclerosis.* Eur J Neurol, 2014. **21**(9): p. 1219-25, e71-2.
23. Platz, T., et al., *Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer Test, Action Research Arm Test and Box and Block Test: a multicentre study.* Clin Rehabil, 2005. **19**(4): p. 404-11.
24. Stark, T., et al., *Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: a systematic review.* PM R, 2011. **3**(5): p. 472-9.
25. Trampisch, U.S., et al., *Optimal Jamar dynamometer handle position to assess maximal isometric hand grip strength in epidemiological studies.* J Hand Surg Am, 2012. **37**(11): p. 2368-73.

26. Paltamaa, J., et al., *Reliability of physical functioning measures in ambulatory subjects with MS.* Physiother Res Int, 2005. **10**(2): p. 93-109.
27. Feys, P., et al., *The Nine-Hole Peg Test as a manual dexterity performance measure for multiple sclerosis.* Mult Scler, 2017. **23**(5): p. 711-720.
28. Verheyden, G., et al., *Reliability and validity of trunk assessment for people with multiple sclerosis.* Phys Ther, 2006. **86**(1): p. 66-76.
29. Chen, C.C. and R.K. Bode, *Psychometric validation of the Manual Ability Measure-36 (MAM-36) in patients with neurologic and musculoskeletal disorders.* Arch Phys Med Rehabil, 2010. **91**(3): p. 414-20.
30. Chen, C.C., et al., *Hand strength and perceived manual ability among patients with multiple sclerosis.* Arch Phys Med Rehabil, 2007. **88**(6): p. 794-7.
31. Lamers, I., et al., *Associations of upper limb disability measures on different levels of the International Classification of Functioning, Disability and Health in people with multiple sclerosis.* Phys Ther, 2015. **95**(1): p. 65-75.
32. Hasson, F., S. Keeney, and H. McKenna, *Research guidelines for the Delphi survey technique.* J Adv Nurs, 2000. **32**(4): p. 1008-15.
33. Landis, J.R. and G.G. Koch, *The measurement of observer agreement for categorical data.* Biometrics, 1977. **33**(1): p. 159-74.
34. Mokkink, L.B., et al., *COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study.* BMC Med Res Methodol, 2020. **20**(1): p. 293.
35. Bertoni, R., et al., *Unilateral and bilateral upper limb dysfunction at body functions, activity and participation levels in people with multiple sclerosis.* Mult Scler, 2015. **21**(12): p. 1566-74.
36. Raats, J., et al., *What is the distribution of trunk impairments and its relationship with disability level in individuals with multiple sclerosis?* Mult Scler Relat Disord, 2022. **57**: p. 103325.
37. Raats, J., et al., *The content and effects of trunk rehabilitation on trunk and upper limb performance in people with multiple sclerosis: a systematic review.* Eur J Phys Rehabil Med, 2022. **58**(1): p. 26-32.
38. Severijns, D., et al., *Motor fatigability after low-intensity hand grip exercises in persons with multiple sclerosis.* Mult Scler Relat Disord, 2016. **10**: p. 7-13.