

A generic outcome assessment of mobility capacity in neurorehabilitation: measurement properties of the de Morton Mobility Index

Tobias Braun (✉ tobias.braun@hs-gesundheit.de)

Hochschule für Gesundheit Bochum <https://orcid.org/0000-0002-8851-2574>

Detlef Marks

Rehaklinik Zihlschlacht <https://orcid.org/0000-0001-8467-3293>

Christian Thiel

Hochschule für Gesundheit Bochum <https://orcid.org/0000-0003-4507-8787>

Christian Grüneberg

Hochschule für Gesundheit Bochum <https://orcid.org/0000-0001-8870-8457>

Research Article

Keywords: neurological rehabilitation, outcome assessment, mobility limitation, reproducibility of results, Rasch analysis, validity

Posted Date: March 18th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-343349/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Mobility capacity is a key outcome domain in neurorehabilitation. The de Morton Mobility Index (DEMMI), an established and generic outcome assessment of mobility in older patients, is promising for use in neurorehabilitation.

Objective: To examine the measurement properties of the DEMMI in rehabilitation inpatients with neurological conditions.

Methods: Cross-sectional study including a mixed sample of adult inpatients in a neurorehabilitation hospital. Structural validity, unidimensionality and measurement invariance (Rasch analysis), construct validity, internal consistency reliability, and inter-rater reliability of the DEMMI (scale range: 0–100 points) were established. The minimal detectable change, the 95% limits of agreement, and possible floor and ceiling effects were calculated to indicate interpretability.

Results: We analyzed validity ($n = 348$) and reliability ($n = 133$) in two samples. In both samples, the majority of participants had a sub-acute stroke or Parkinson's disease.

Rasch analysis indicated unidimensionality with an overall fit to the model ($\chi^2 = 59.4$, $P = 0.074$). There was no relevant measurement invariance by disease group. All 11 hypotheses on construct validity were confirmed. Internal consistency reliability (Cronbach's $\alpha = 0.94$) and inter-rater reliability (intraclass correlation coefficient = 0.94; 95% confidence interval: 0.91–0.95) were sufficient. The minimal detectable change with 90% confidence was 15.0 points and the limits of agreement were 39%. No floor or ceiling effects were observed.

Conclusions: Results indicate sufficient measurement properties of the DEMMI in rehabilitation inpatients with neurological conditions. The DEMMI can be used as a generic outcome assessment of mobility capacity in neurorehabilitation.

Trial registration: German Clinical Trials Register (DRKS00004681). Registered May 6, 2013.

Introduction

For individuals with neurological conditions, mobility limitations are a frequent and significant issue which negatively affect independence in daily living and quality of life, and increase the risk of falls.^{1,2} Thus, improvements in mobility capacity, especially in walking and balancing, are considered the most important rehabilitation goals of patients with neurological conditions.³⁻⁵

Although guideline-directed interventions vary for different neurological conditions, such as stroke or Parkinson's disease (PD),^{6,7} patients are often treated in inpatient or outpatient rehabilitation facilities which are not focused on a single disorder. In such settings, generic outcome assessments are used to

measure outcome domains that are common in many disease groups. For example, the Functional Independence Measure is used as generic measure of disability across different neurological conditions.^{8,9}

Mobility capacity, too, is a health-related outcome domain relevant to neurorehabilitation that is often assessed across different disease groups with generic outcome assessments. However, commonly used assessments, such as timed walking tests for gait speed, the timed up and go test, and the 6-minute walk test,^{5,10-12} cover only a limited range of mobility components (e.g., walking on plane surface) and their clinical usability is limited, because they require the patient's ability to walk. Thus, a considerable number of individuals cannot be assessed, particularly in the early and sub-acute stages of rehabilitation after an acute neurological event, such as a stroke.^{13,14} Ceiling effects also impact clinical utility, particularly when an assessment is intended to measure progress over the duration of recovery.¹⁰

An ideal generic outcome assessment of mobility capacity in neurorehabilitation needs to fulfil the following characteristics: performance-based, measure on interval level, affordable, easy to learn for assessors, feasible, safe, valid over the whole mobility spectrum, sound measurement properties (e.g. validity, reliability, responsiveness), and invariant across disease groups and other patient characteristics.¹² A clinical outcome assessment that fulfils these requirements in geriatric care is the de Morton Mobility Index (DEMMI)¹⁵. The DEMMI was developed based on the Rasch model^{16,17} to measure the mobility capacity of older hospital patients¹⁵ – a heterogenous population of individuals with a broad range of limitations in self-care, mobility, and cognition. This performance-based clinical outcome assessment has a broad scale width, covering low- to high-order mobility abilities and producing interval-level test scores. The DEMMI's measurement properties have been examined in various health conditions and care settings, indicating sufficient validity, reliability, responsiveness to change and unidimensionality in acute and sub-acute older patients,^{15,18-20} osteoarthritis,²¹ hip fracture,²² dementia,²³⁻²⁵ and critically ill patients.²⁶ Some studies provide very promising evidence that the DEMMI is a feasible, valid, reliable and unidimensional assessment for individuals with neurological conditions, such as stroke,^{27,28} PD,^{29,30} and mixed neurological conditions.²⁷ The DEMMI form fits on one paper sheet and can be administered by health professionals within 5–10 minutes without special equipment.^{20,23,15}

We hypothesized that the DEMMI is a generic outcome assessment of mobility capacity in neurorehabilitation. Thus, we aimed to evaluate the DEMMI's measurement properties in a mixed sample of rehabilitation inpatients with neurological conditions.

Methods

Reporting of this study was informed by the STROBE guideline for observational studies, the GRRAS guideline for reliability studies and the criteria of the COSMIN risk of bias checklist.³¹⁻³⁴

Design and Setting

We performed a cross-sectional study of the DEMMI's measurement properties in neurorehabilitation. This study was approved by the Local Committee for Ethics in Medical Research (Canton of Thurgau, Switzerland: 2013/13), was registered *a priori* (German Clinical Trials Register: DRKS00004681), and all participants gave written informed consent. Briefly, rehabilitation inpatients with neurological conditions were examined with the DEMMI and a set of functional assessments (listed below) on several occasions to analyze the DEMMI's psychometric properties. The present study reports on the DEMMI's structural and construct validity, internal consistency, inter-rater reliability, measurement error, interpretability, and feasibility for the complete sample of rehabilitation inpatients with neurological conditions. The DEMMI's measurement properties for sub-samples with stroke and PD have been published previously.^{28,29}

The study was conducted in a neurological rehabilitation hospital in Switzerland, where patients were typically referred from acute hospitals, neurologist consultants, or general practitioners located in the eastern and central parts of Switzerland.

Participants

The study sample consisted of all inpatients present on May 8, 2013 or entering the rehabilitation hospital consecutively within the following 20 weeks. Inclusion criteria were a neurological disorder and an age of 18 years and older. The main exclusion criteria were severe cognitive impairment and a contraindication for mobilization (for all criteria, see Figure 1).

Procedures

Eligible participants were examined by the primary investigator (TB) in a single session of 30–45 minutes scheduled within the first 7 days after hospital admission, if possible. The DEMMI and a comprehensive set of functional assessments were performed in a standardized order (baseline).

The participants' socio-demographic data were taken from the medical records. For common disorders, disease-specific measures were performed to describe disease severity and functional capacity. For participants with stroke, the National Institutes of Health Stroke Scale was assessed to measure the global severity of stroke symptoms.³⁵ For participants with PD and Multiple Sclerosis, Hoehn and Yahr staging³⁶ and the Expanded Disability Status Scale³⁷ were completed by the hospital neurologist, respectively. In all three scales, higher scores indicate higher impairment or disease severity.

Inter-rater reliability was examined between 2 trained and experienced physiotherapists, the primary investigator (TB) and a second rater (DM). Characteristics of both raters are described elsewhere.^{28,29}

The second rater performed the DEMMI independently in a convenient sub-sample (reliability sample). Participant selection was mainly based on the second rater's availability (temporal resources) and on participants' consent to perform a second study assessment. Both DEMMI assessments were performed within 2 days. To create a stable retest situation, participants were excluded if they reported a change in their physical or mental condition with respect to the first session (e.g., fatigue, pain, ON/OFF state in PD).

The test environment (patient's room) was similar for both sessions (baseline and retest). Both raters were blinded toward each other's ratings and we tried to balance the number of participants each rater visited first.

A sample size of ≥ 50 participants for reliability studies has been proposed to be "good" at the times of study conduction.^{38,39} However, within the initial recruitment period (20 weeks), we could not include ≥ 50 participants for each major sub-sample of participants with stroke and PD, respectively. Hence, we set up a second recruitment period, using the same inclusion criteria, and screened all present and incoming patients over a period of 9 consecutive days. This additional sample of convenience was only included in the inter-rater reliability analysis.

Measurements

Participants were assessed with the DEMMI, together with a set of functional assessments, including Berg Balance Scale, timed up and go test, 10-meter walk test, Functional Ambulation Categories (FAC), 6-minute walk test, Performance Oriented Mobility Assessment, and Functional Independence Measure. For the sub-samples of participants with stroke and PD, we performed additional functional assessments, which were only used to analyze these sub-samples.^{28,29}

A detailed description of the assessment procedures and a description of the comparator assessments are given in the Supplementary file 1. Table 1 provides an overview of the scale width and constructs measured by the comparator assessments.

DEMMI

The DEMMI is a performance-based clinical outcome assessment of mobility capacity, consisting of 15 hierarchical mobility items.^{15,40,20,41} The patient is asked to perform functional tasks related to bed and chair mobility, ambulation, static balance, and dynamic balance. The items are rated with 2-or 3-point response options, resulting in a maximum ordinal score of 19 points. This raw score is transformed into a total interval DEMMI score of 0–100 points, with higher scores indicating a higher level of mobility capacity.

Statistical Analysis

Data were analyzed using SPSS version 23.0 and Microsoft Excel (Professional Plus 2016) for all analyses except the Rasch analysis, which was completed using RUMM2030 version 5.1 software. Descriptive statistics were used to present sample characteristics. Interval-based data were examined for normal distribution with the Shapiro-Wilk test of normality and by visual inspection of the related histograms and P-P-plots. The DEMMI scores were not normally distributed ($p < 0.001$); therefore, only non-parametric statistics were applied. A significance value of 5% was used.

Measurement properties

Structural validity (Rasch analysis)

The Rasch model is a probabilistic model asserting that item response is a logistic function of item difficulty and person ability¹⁶ The DEMMI was developed based on the Rasch model in geriatric inpatients¹⁵ and data fitted the model in various other medical conditions.^{22,20,23,28,30}

We performed a Rasch analysis to evaluate the following properties of the DEMMI in neurological inpatients: stochastic (probabilistic) ordering of items, monotonicity (increase in item responses consistent with the underlying trait), local item independence (zero correlation between items when conditioned on the score), unidimensionality, and group invariance (no difference in response to item by group membership when at the same level of (in this case) 'mobility capacity'), which is also called differential item functioning (DIF). Data fit to the model was deemed acceptable if a set of criteria was fulfilled (Supplementary file 2). Full details of the Rasch analysis process are given elsewhere.^{42,17} Reporting followed established recommendations.¹⁷

A target sample size of at least 150 was set to provide 99% confidence within ± 0.5 logits.⁴³ The unrestricted (partial credit) Rasch polytomous model was used with a conditional pair-wise parameter estimation.

Construct validity

In absence of a 'gold standard' for 'mobility capacity', construct validity was assessed by following the methodological approach of hypotheses testing.^{38,39} We used the other functional outcomes and participants' clinical information to assess the DEMMI's construct validity. Aspects of convergent and known-groups validity were used to formulate 11 hypotheses (H1–H15).^{39,44} All hypotheses were formulated *a priori*, based on existing literature, and the clinical expertise of clinicians and the research team.^{30,15,45,23,20} Formulated and shortened versions of the hypotheses are presented in Supplementary file 1 and Table 1, respectively. Details on the statistical analyses and interpretation of hypotheses testing are given in Supplementary file 1. A sample size of ≥ 100 participants is recommended.⁴⁶

Reliability

Cronbach's alpha and the Person-Item-Separation Index, which are measures of internal consistency reliability in case of a unidimensional scale, were derived from the validity sample because of its larger sample size.³⁹ An outcome between 0.70 and 0.95 was considered acceptable.³⁹

Inter-rater reliability was examined using the intra-class correlation coefficient (ICC) model 2.1 (two-way random effects model; ICC_{AGREEMENT}).⁴⁴ An ICC of ≥ 0.7 or higher was deemed acceptable.³⁹ The standard error of measurement (SEM_{AGREEMENT}) was calculated and deemed satisfactory if it was $\leq 10\%$ of the total scale range (100 DEMMI points).^{47,44} The absolute and relative agreement between both raters per DEMMI item was calculated as a percentage (%) and as the weighted kappa with linear weights

(κ).⁴⁴ Agreement per item equal or above 70% and $\kappa \geq 0.70$ was considered acceptable.³⁹ For additional information on reliability statistics, see Supplementary file 1.

Interpretability

Bland and Altman's method was used to illustrate agreement between the two raters.⁴⁸ The minimal detectable change (MDC) with 90% and 95% confidence was calculated for individual subjects (MDC_{ind}) as well as for comparisons of mean scores between groups (MDC_{group}).^{49,44} A floor or ceiling effect was considered if $\geq 15\%$ of the participants scored the highest or lowest possible DEMMI score.³⁹ Supplementary file 1 gives more information on the statistical methods.

Feasibility

We calculated the mean administration time for the DEMMI in minutes and related the administration time to the participants' functional status. We documented any adverse events, such as falls, reports of pain, atypical and severe changes of muscle tone, or significant fatigue.

Results

Of 505 neurological inpatients screened for eligibility, 348 (69%) were assessed within the first recruitment period for the validity sample. For the inter-rater reliability analysis ($n = 133$), 102 participants could be reassessed and an additional sample of 31 participants was recruited. Figure 1 shows the flow of participants throughout the study. Table 2 gives the participants' demographics according to psychometric sampling.

The participants' mean age was 66 ± 13 years, 218 (63%) were male, and 230 (66%) were able to ambulate independently in the hospital (FAC level ≥ 4). Most participants had a stroke (nontraumatic intracerebral hemorrhage or cerebral infarction; $n = 126$; 36%), including 109 (31%) in the sub-acute phase (< 6 months after stroke onset) and 17 (5%) in the chronic phase (≥ 6 months). Of 108 (31%) participants with an extrapyramidal or movement disorder, 100 (29%) had PD. Other frequent disorders were multiple sclerosis ($n = 18$; 5%), neoplasms of the brain or the central nervous system ($n = 18$; 5%) and traumatic brain injury ($n = 13$; 4%). Fifteen (4%) participants presented with a non-traumatic spinal cord injury due to various diseases, such as intervertebral disc disorders. The detailed sample composition according to ICD-10 diagnoses is presented in the table in Supplementary file 2.

In 78% of the participants, the study assessment was performed within the first 7 days after admission. There were no missing items for any DEMMI assessment. The distribution of DEMMI scores is illustrated in the figure in Supplementary file 2. Table 1 includes the mobility-related outcomes for all comparator assessments.

Structural validity (Rasch analysis)

Rasch analysis was performed on the complete DEMMI item sets of 348 participants and on the complete 15-item scale. Summary fit statistics are given in Table 3. There was overall fit to the model, including no mis-fitting persons and no mis-fitting items. We found no disordered thresholds, indicating that the responses to the items were consistent with the metric estimate of the underlying construct of mobility capacity. Unidimensionality was confirmed and data were free of local dependency. Overall, the participants exhibited a higher level of mobility (mean: 2.1 logits) than the scale average (0.0 logits; person-item distribution map in Supplementary file 2).

There was no DIF (measurement invariance) by sex, age, or disease phase (sub-acute and chronic). There was uniform DIF by disease group (stroke, PD, and 'other') for two items. Participants with PD were less likely to achieve item #2 'roll' (found it harder; $F = 7.8$, $p < 0.001$) and more likely to achieve item #15 'jump' (found it easier; $F = 9.7$, $p < 0.001$) than participants in the other two diagnosis groups, respectively (figures of the Item Characteristic Curves in Supplementary file 2). Further assessment indicated that both items showed 'real' (and no artificial) DIF.⁵⁰

The importance of DIF exhibited by item #2 and #15 was further assessed, as recommended,⁵¹ by comparing the Rasch estimate between a 'pure' dataset (excluding these 2 items) and a fully anchored dataset. There were trivial, non-significant differences in mean individual person logits between the two datasets (without DIF items mean: 2.34 ± 4.55 , with DIF items mean: 2.26 ± 3.94 , $p = 0.076$). In total, 33.0% of the person estimates differed by more than 0.5 logits, but the correlation between the two sets of person estimates was quite high (ICC = 0.981, 95% CI: 0.976–0.984). Overall, the results indicate that the identified DIF was unimportant.

The DEMMI's item hierarchy in the sample of inpatients in neurorehabilitation compared to that of the development sample of geriatric inpatients¹⁵ is illustrated in the figure in Supplementary file 2. A high positive logit location (e.g., tandem standing with eyes closed) indicates harder item difficulty compared to a negative logit location (e.g., sit unsupported). Deviations from the original item hierarchy are indicated by non-overlapping 95% confidence bands in 6 items.

Construct validity

All 11 (100%) a priori stated hypotheses about correlations of the DEMMI with other clinical outcome assessments and known-group differences were confirmed. Correlations between the DEMMI and other broad outcome measures of mobility, ambulation, walking endurance, balance and functional independence (H1–H8; Table 1) were between 0.73 and 0.94. Groups of participants who walked without a walking aid (H9), who were independent in walking (H10), and who were able to climb stairs (H11) had significantly higher DEMMI mean scores than the less able comparison groups (Supplementary file 2).

Reliability

Cronbach's alpha and the Person-item Separation Index of the DEMMI were 0.94 and 0.90, respectively, indicating excellent internal consistency reliability.

For the inter-rater reliability analysis, the DEMMI assessment was performed twice by two different physiotherapists on 133 participants. Sample characteristics are given in Table 2. Rater 1 (TB) administered the first DEMMI measure in 85 (64%) participants. The two DEMMI assessments were performed on the same day in 77 (58%) participants, and within 2 days in 56 (42%) participants. There was no statistically significant mean difference in DEMMI scores between both assessors (0.1 points; 95% CI: -1.5–1.6; $P = 0.92$) and there was no considerable variance (0.3) due to systematic differences between the two raters. The variance between participants was 616.7 and the residual variance was 42.3. The $ICC_{AGREEMENT}$ was 0.94 (95% CI: 0.91–0.95).

Measurement error

The $SEM_{AGREEMENT}$ was 6.5 points and considered acceptable (6.5% of the total DEMMI scale range).

The absolute and relative agreement per item are presented in the table in Supplementary file 2. There was no DEMMI item with absolute agreement < 70% (range: 80% to 99%), but 3 items with $\kappa < 0.7$ (range: 0.29 to 0.92).

Interpretability

The Bland-Altman plot is illustrated in Figure 2. The data were heteroscedastic ($\tau = 0.26$) and differences were not normally distributed ($P < 0.01$). The 95% limits of agreement were $0.39X + 0.1$ and $-0.39X + 0.1$, respectively, with X denoting the mean score.

The MDC_{ind90} , MDC_{ind95} , $MDC_{group90}$, and $MDC_{group95}$ were 15.0, 18.0, 1.3, and 1.6 points, respectively.

There were no absolute floor or ceiling effects, with 15 (4%) participants scoring 0 and 21 (6%) participant scoring 100 DEMMI points, respectively (histogram in Supplementary file 2).

Feasibility

The mean administration time of 100 DEMMI assessments was 6.3 ± 2.1 (range: 1–14) minutes (figure in Supplementary file 2). DEMMI administration took ≤ 10 minutes in 96% ($n = 333$) of participants. In non-ambulant or dependent walkers ($FAC \leq 3$, $n = 118$) and independent walkers ($FAC \geq 4$, $n = 230$), the administration time was 6.8 ± 2.7 and 6.0 ± 1.6 minutes, respectively. No adverse events occurred in any DEMMI assessment.

Discussion

This study provides evidence of the DEMMI's sound structural and construct validity, internal consistency, inter-rater reliability, measurement error, interpretability, and feasibility in a mixed sample of rehabilitation inpatients with neurological conditions.

Rasch analysis confirmed structural validity in terms of unidimensionality, hierarchical order, measurement invariance, and logistic item structure. This is in keeping with results of other studies supporting the DEMMI's structural validity examined in geriatric^{15,22,20,23} and neurological populations.^{30,28,29} Two facets seem notably important because of their clinical relevance. Unidimensionality indicates that the DEMMI measures one single underlying construct (mobility capacity). Measurement invariance (no DIF) indicates that DEMMI items do not function differently for different members of a sample group (e.g., for men and women or for different disease groups). No (important) DIF was observed by sex, age, disease phase or disease group; therefore, clinicians and researchers can be confident that the DEMMI is a 'fair' test. That means, that every patient/person with the same level of mobility capacity has the same change to endorse each DEMMI items and exceed the same final DEMMI score.

The DEMMI's construct validity in neurorehabilitation is indicated by strong correlations with other commonly applied and validated outcome assessments of mobility, ambulation, walking endurance, balance, and functional independence. The following results strengthen the conclusion of sufficient construct validity: All hypotheses were confirmed; even the lower confidence bounds of most correlations were > 0.7 ; and the difference in DEMMI mean scores between clinical groups was larger than the minimal important change of 10 DEMMI points.¹⁵ This result is not surprising, since strong correlations with other measures of mobility have been reported consistently in older patients and individuals in neurorehabilitation.^{15,28,29,27}

Cronbach's alpha (0.94) was within the proposed acceptable range of 0.70 to 0.95 and can be judged as excellent.³⁹ The inter-rater reliability of the DEMMI performed by two experienced physiotherapists was 0.94 (95% CI: 0.91–0.95) and is comparable to other inter-rater reliability estimations between 0.85 and 0.94 reported by others.^{15,18,21,23} An ICC of ≥ 0.7 is considered sufficient for group comparisons, and a value of ≥ 0.90 is an indicator of acceptable reliability for individual-level monitoring.^{39,52}

Although the ICC is quite high, the DEMMI is not free of measurement error. The SEM (6.5 points) is considered acceptable and comparable to other estimations (4.1 to 7.5 points).^{20,28,29} In addition, there was no item with absolute agreement $< 70\%$. This evidence of sufficient inter-rater reliability has some crucial clinical implications in those clinical situations, in which the DEMMI is assessed on a single patient twice by two different assessors (e.g., different physiotherapists on admission and discharge). In this situation, provided that both assessors carefully synchronize before clinical use, one can be very confident that each DEMMI value represents the 'true' level of mobility capacity of that individual patient at that time, and that the different assessors would obtain similar scores.

We used the reliability data to establish information on the DEMMI's interpretability.³⁸ We found relatively large limits of agreement (39%), which were in line with previous values reported for patients with PD (31%),²⁹ sub-acute stroke (42%),²⁸ and sub-acute geriatric conditions (-8.4 to 11.8 points).²⁰ The MDC_{ind90} value of 15.0 points is considerably higher than the MDC_{ind} range of 6 to 10 points reported for

older adults,^{19,18,15,20,53} but comparable to the MDC_{ind} values of 12.5 and 17.5 reported for the sub-samples with stroke and PD, respectively.^{29,28} Thus, a DEMMI change score, assessed by two different assessors at two different time points, needs to be ≥ 15 points (or $\pm 39\%$) to have high confidence that this change score is free of measurement error.

A possible explanation for this relatively large MDC value (and limits of agreement) could be the high variability of mobility capacity in the reliability sample (standard deviation of 25 points = 42% of the sample's mean score). A further explanation could be that the calculation of MDC values performed here include the inter-rater variance and the participants' intra-individual variance. Thus, the comparably large MDC values might be biased and overestimated by the included inter-rater variance. Our MDC estimations should be considered with caution and verified by future studies, which should use test-retest reliability estimations for stable patients generated by a single assessor.^{49,44}

This study provides evidence for the DEMMI's high feasibility over the whole mobility spectrum of individuals with neurological conditions, since no floor or ceiling effects occurred at hospital admission. Other authors also reported no floor or ceiling effects on admission in samples of patients with PD and stroke.^{28-30,27} This might be an important advantage of the DEMMI over other established outcome assessments of mobility and ambulation in neurorehabilitation. We observed significant floor effects (approximately 20%) for the Timed Up and Go test, gait speed assessment, and the 6-minute walk test (Table 2). With these assessments, longitudinal monitoring of mobility capacity from admission to a later point in rehabilitation would not have been possible. Similar floor effects of gait assessments in patients with stroke have been reported.^{13,14} However, a mild ceiling effect (19%) for the DEMMI in patients with stroke at hospital discharge has been reported in one study.²⁷ In the present sample, only 6% of participants scored the highest DEMMI score of 100 points at hospital admission. The DEMMI includes some high-level mobility items (e.g., jump, tandem standing with eyes closed) and further research is needed to evaluate whether patients who complete these items successfully (and reach the maximum score) suffer from subjective or objective mobility limitations at all; especially compared to healthy individuals of the same age.

The mean administration time of 6.3 minutes achieved by an experienced assessor is comparable to previous findings.^{15,20,23,28,29} Based on existing evidence and our clinical experience, the DEMMI can be completed within 5–10 minutes in most neurological inpatients by a trained healthcare professional. High feasibility, information on interpretability, and short administration times of outcome assessments facilitate routine clinical application and enlarge therapy time.⁵⁴

Strengths and limitations

We examined a broad set of measurement properties in a sufficiently large^{46,43} and consecutive sample of hospital inpatients with neurological conditions, supporting the generalizability of results. The included participants presented with a wide spectrum of disability, age range (18–90 years), disease duration, sub-

acute and chronic conditions, and various diseases. However, the external validity of this study might be limited because the data were collected in a single rehabilitation hospital only and, with respect to prevalence estimations for central Europe, participants with PD were over-represented, whereas other conditions, such as Multiple Sclerosis and spinal cord injury, were underrepresented.⁵⁵⁻⁵⁷

We used a combination of modern methods of latent trait theory (Rasch analysis) and methods of classical test theory. Rasch analysis provides an especially powerful tool to analyze unidimensionality, measurement invariance, and logistic item structure.¹⁷

We did not evaluate the DEMMI's floor and ceiling effects at any later time of rehabilitation (e.g., hospital discharge). As previously mentioned, MDC values and limits of agreement need to be interpreted with caution and might be lower than reported in this study for test-retest situations.

Implications for clinical practice and further research

This study provides evidence that the DEMMI has sufficient key measurement properties in neurorehabilitation, including structural validity and internal consistency reliability. It seems feasible and safe, since no adverse events occurred during or immediately after test administration, which took only 6 minutes on average. The lack of any floor or ceiling effects on hospital admission indicates clinical value and applicability across the whole mobility spectrum of inpatients with neurological conditions. For the DEMMI administration, no long training period is required, no special equipment is needed, and there is no license charge. These advantages address some barriers to the use of outcome assessments by healthcare professionals,^{54,58} and they could facilitate the application of this instrument in clinical care.

We found no relevant DIF by disease phase and disease group, indicating that the DEMMI measures were free of measurement invariance in this sample and that the DEMMI can be used as a generic measure of mobility capacity in this population.

Further research should focus on measurement properties that are still unknown in neurorehabilitation, such as test-retest reliability, responsiveness, minimal important change values, and prognostic validity. Replication of our findings is warranted⁵⁹ and should also evaluate the DEMMI's psychometric properties in mixed samples with other compositions, including more individuals of other disease groups.

Since there are many assessments available to measure mobility capacity in neurorehabilitation, the DEMMI's psychometric quality and clinical utility should be compared to other (generic) assessments in clinical trials and systematic reviews that follow recommended methods.⁶⁰ Future studies should investigate the DEMMI's clinical utility for goal setting and guiding rehabilitation strategies.

Conclusions

The DEMMI seems to be a unidimensional, valid, and reliable performance-based clinical outcome assessment of mobility capacity in adult individuals with neurological conditions that can be used generically in this population. Provided that the high feasibility, clinical utility, and sufficient measurement properties found in this study are confirmed in future studies, the DEMMI might become the standard assessment of mobility capacity in neurorehabilitation.

Declarations

Acknowledgments

We thank all participants for their time and contributions. We further acknowledge the support of the physiotherapy and medical staff of the Rehaklinik Zihlschlacht.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Authors' contributions

Study concept and design: TB, DM, CG. Acquisition of data: TB, DM. Analysis of data: TB. Interpretation of data: TB, DM, CT, CG. Drafting the manuscript: TB. Manuscript revision for important intellectual content: DM, CT, CG. All authors read and approved the final version of the manuscript.

Funding:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

ORCID ID

Tobias Braun: <https://orcid.org/0000-0002-8851-2574>

Detlef Marks: <https://orcid.org/0000-0001-8467-3293>

Christian Thiel: <https://orcid.org/0000-0003-4507-8787>

Christian Grüneberg: <https://orcid.org/0000-0001-8870-8457>

References

1. World Health Organization. *International Classification of Functioning, Disability and Health: ICF*. World Health Organization; 2001.

2. Mayo NE, Wood-Dauphinee S, Côté R, Durcan L, Carlton J. Activity, participation, and quality of life 6 months poststroke. *Arch Phys Med Rehabil*. 2002;83(8):1035-1042.
3. Rice DB, McIntyre A, Mirkowski M, et al. Patient-Centered Goal Setting in a Hospital-Based Outpatient Stroke Rehabilitation Center. *PM R*. 2017;9(9):856-865. doi:10.1016/j.pmrj.2016.12.004
4. Plant S, Tyson SF. A multicentre study of how goal-setting is practised during inpatient stroke rehabilitation. *Clin Rehabil*. 2018;32(2):263-272. doi:10.1177/0269215517719485
5. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. *Lancet*. 2011;377(9778):1693-1702. doi:10.1016/S0140-6736(11)60325-5
6. Keus S, Munneke M, Graziano M, et al. *European Physiotherapy Guideline for Parkinson's Disease*. KNGF/ParkinsonNet; 2014.
7. Winstein CJ, Stein J, Arena R, et al. Guidelines for Adult Stroke Rehabilitation and Recovery: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association. *Stroke*. 2016;47(6):e98-e169. doi:10.1161/STR.0000000000000098
8. Khoo TC, FitzGerald A, MacDonald E, Bradley L. Outcomes for older adults in inpatient specialist neurorehabilitation. *Ann Phys Rehabil Med*. 2020;63(4):340-343. doi:10.1016/j.rehab.2019.05.001
9. Ottenbacher KJ, Hsu Y, Granger CV, Fiedler RC. The reliability of the functional independence measure: a quantitative review. *Arch Phys Med Rehabil*. 1996;77(12):1226-1232.
10. Ashford S, Brown S, Turner-Stokes L. Systematic review of patient-reported outcome measures for functional performance in the lower limb. *J Rehabil Med*. 2015;47(1):9-17. doi:10.2340/16501977-1889
11. Pollock C, Eng J, Garland S. Clinical measurement of walking balance in people post stroke: a systematic review. *Clin Rehabil*. 2011;25(8):693-708. doi:10.1177/0269215510397394
12. Hobart JC, Cano SJ, Zajicek JP, Thompson AJ. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations. *Lancet Neurol*. 2007;6(12):1094-1105. doi:10.1016/S1474-4422(07)70290-9
13. Jørgensen HS, Nakayama H, Raaschou HO, Olsen TS. Recovery of walking function in stroke patients: the Copenhagen Stroke Study. *Arch Phys Med Rehabil*. 1995;76(1):27-32.
14. Hafsteinsdottir TB, Rensink M, Schuurmans M. Clinimetric properties of the Timed Up and Go Test for patients with stroke: a systematic review. *Top Stroke Rehabil*. 2014;21(3):197-210. doi:10.1310/tsr2103-197
15. de Morton NA, Davidson M, Keating JL. The de Morton Mobility Index (DEMMI): an essential health index for an ageing world. *Health Qual Life Outcomes*. 2008;6:63. doi:10.1186/1477-7525-6-63
16. Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests*. Expanded ed. University of Chicago Press; 1980.
17. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum*. 2007;57(8):1358-1362. doi:10.1002/art.23108

18. de Morton N, Davidson M, Keating JL. Reliability of the de Morton mobility index (DEMMI) in an older acute medical population. *Physiother Res Int*. 2010;16:159-169. doi:10.1002/pri.493
19. de Morton NA, Lane K. Validity and reliability of the de Morton Mobility Index in the subacute hospital setting in a geriatric evaluation and management population. *J Rehabil Med*. 2010;42(10):956-961. doi:10.2340/16501977-0626
20. Braun T, Schulz R-J, Reinke J, et al. Reliability and validity of the German translation of the de Morton Mobility Index (DEMMI) performed by physiotherapists in patients admitted to a sub-acute inpatient geriatric rehabilitation hospital. *BMC Geriatr*. 2015;15(1):1660. doi:10.1186/s12877-015-0035-y
21. Jans MP, Slootweg VC, Boot CR, de Morton NA, van der Sluis G, van Meeteren NL. Reproducibility and validity of the Dutch translation of the de Morton Mobility Index (DEMMI) used by physiotherapists in older patients with knee or hip osteoarthritis. *Arch Phys Med Rehabil*. 2011;92(11):1892-1899. doi:10.1016/j.apmr.2011.05.011
22. de Morton NA, Harding KE, Taylor NF, Harrison G. Validity of the de Morton Mobility Index (DEMMI) for measuring the mobility of patients with hip fracture during rehabilitation. *Disabil Rehabil*. 2013;35(2):105-111. doi:10.3109/09638288.2012.705220
23. Braun T, Grüneberg C, Thiel C, Schulz R-J. Measuring mobility in older hospital patients with cognitive impairment using the de Morton Mobility Index. *BMC Geriatr*. 2018;18:100. doi:10.1186/s12877-018-0780-9
24. Braun T, Schulz R-J, Thiel C, Grüneberg C. Test-retest reliability of the de Morton Mobility Index (DEMMI) in older acute medical patients with cognitive impairment (Poster); July 2, 2017; Cape Town, South Africa.
25. Braun T, Thiel C, Schulz R-J, Grüneberg C. Responsiveness and interpretability of commonly used outcome assessments of mobility capacity in older hospital patients with cognitive spectrum disorders. *Health Qual Life Outcomes*. 2021;19(1):68. doi:10.1186/s12955-021-01690-3
26. Sommers J, Vredeveld T, Lindeboom R, Nollet F, Engelbert RH, van der Schaaf M. The de Morton Mobility Index Is Feasible, Reliable, and Valid in Critically Ill Patients. *Phys Ther*. 2016;96(10):1658-1666. doi:10.2522/ptj.20150339
27. New PW, Scroggie GD, Williams CM. The validity, reliability, responsiveness and minimal clinically important difference of the de Morton mobility index in rehabilitation. *Disabil Rehabil*. 2017;39(10):1039-1043. doi:10.1080/09638288.2016.1179800
28. Braun T, Marks D, Thiel C, Grüneberg C. Reliability and validity of the de Morton Mobility Index in individuals with sub-acute stroke. *Disabil Rehabil*. 2019;41(13):1561-1570. doi:10.1080/09638288.2018.1430176
29. Braun T, Marks D, Thiel C, Menig A, Grüneberg C. An investigation of the measurement properties of the de Morton Mobility Index for measuring mobility capacity in hospital patients with Parkinson's disease. *Clin Rehabil*. 2021;35(3):423 –435. doi:10.1177/0269215520966472
30. Johnston M, de Morton N, Harding K, Taylor N. Measuring mobility in patients living in the community with Parkinson disease. *NeuroRehabilitation*. 2013;32(4):957-966. doi:10.3233/NRE-

31. Elm E von, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*. 2008;61(4):344-349. doi:10.1016/j.jclinepi.2007.11.008
32. Kottner J, Audige L, Brorson S, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol*. 2011;64(1):96-106. doi:10.1016/j.jclinepi.2010.03.002
33. Mokkink LB, Vet HCW de, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res*. 2018;27(5):1171-1179. doi:10.1007/s11136-017-1765-4
34. Mokkink LB, Boers M, van der Vleuten CPM, 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):293. doi:10.1186/s12874-020-01179-5
35. Criddle LM, Bonnono C, Fisher SK. Standardizing stroke assessment using the National Institutes of Health Stroke Scale. *J Emerg Nurs*. 2003;29(6):541-546. doi:10.1016/j.jen.2003.08.011
36. Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. *Neurology*. 1967;17(5):427-442.
37. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*. 1983;33(11):1444-1452. doi:10.1212/wnl.33.11.1444
38. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737-745. doi:10.1016/j.jclinepi.2010.02.006
39. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42. doi:10.1016/j.jclinepi.2006.03.012
40. Braun T, Schulz R-J, Hoffmann M, et al. German version of the de Morton Mobility Index. First clinical results from the process of the cross-cultural adaptation. *Z Gerontol Geriatr*. 2015;48(2):154-163. doi:10.1007/s00391-014-0648-3
41. Braun T, Rieckmann A, Grüneberg C, Marks D, Thiel C. Hierarchical assessment of balance and mobility - German translation and cross-cultural adaptation. *Zeitschrift für Gerontologie und Geriatrie*. 2016;49(5):386-397. doi:10.1007/s00391-016-1026-0
42. Pallant JF, Tennant A. An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *Br J Clin Psychol*. 2007;46(Pt 1):1-18.
43. Linacre JM. Sample size and item calibration stability. *Rasch Measurement Transactions*. 1994;7(4):328.
44. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide*. Cambridge University Press; 2011.
45. Braun T, Grüneberg C, Coppers A, Tofaute L, Thiel C. Comparison of the de Morton Mobility Index and Hierarchical Assessment of Balance and Mobility in older acute medical patients. *J Rehabil Med*.

2018;50:292-301. doi:10.2340/16501977-2320

46. Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, Vet HCW de. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res.* 2012;21(4):651-657. doi:10.1007/s11136-011-9960-1
47. van Bloemendaal M, Bout W, Bus SA, Nollet F, Geurts AC, Beelen A. Validity and reproducibility of the Functional Gait Assessment in persons after stroke. *Clinical Rehabilitation.* 2019;33(1):94-103. doi:10.1177/0269215518791000
48. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1(8476):307-310.
49. Stratford PW, Binkley JM, Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. *Phys Ther.* 1996;76(10):1109-1123.
50. Andrich D, Hagquist C. Real and Artificial Differential Item Functioning. *Journal of Educational and Behavioral Statistics.* 2012;37(3):387-416. doi:10.3102/1076998611411913
51. Tennant A, Pallant JF. DIF matters: A practical approach to test if differential item functioning makes a difference. *Rasch Measurement Transactions.* 2007;20(4):1082-1084.
52. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res.* 2002;11(3):193-205.
53. Braun T, Thiel C, Schulz R-J, Grüneberg C. Reliability of mobility measures in older medical patients with cognitive impairment. *BMC Geriatr.* 2019;19:20. doi:10.1186/s12877-019-1036-z
54. Braun T, Rieckmann A, Weber F, Grüneberg C. Current use of measurement instruments by physiotherapists working in Germany: A cross-sectional online survey. *BMC Health Serv Res.* 2018;18:810. doi:10.1186/s12913-018-3563-2
55. Kingwell E, Marriott JJ, Jetté N, et al. Incidence and prevalence of multiple sclerosis in Europe: a systematic review. *BMC Neurol.* 2013;13(1):128. doi:10.1186/1471-2377-13-128
56. Wyndaele M, Wyndaele J-J. Incidence, prevalence and epidemiology of spinal cord injury: what learns a worldwide literature survey? *Spinal Cord.* 2006;523-529. doi:10.1038/sj.sc.3101893
57. Cattaneo D, Gervasoni E, Pupillo E, et al. Mobility Disorders in Stroke, Parkinson Disease, and Multiple Sclerosis: A Multicenter Cross-Sectional Study. *Am J Phys Med Rehabil.* 2020;99(1):41-47. doi:10.1097/PHM.0000000000001272
58. van Peppen RPS, Maissan FJF, van Genderen FR, van Dolder R, van Meeteren NLU. Outcome measures in physiotherapy management of patients with stroke: a survey into self-reported use, and barriers to and facilitators for use. *Physiother Res Int.* 2008;13(4):255-270. doi:10.1002/pri.417
59. Ioannidis JPA. Why replication has more scientific value than original discovery. *The Behavioral and brain sciences.* 2018;41:e137. doi:10.1017/S0140525X18000729
60. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res.* 2018;27(5):1147-1157. doi:10.1007/s11136-018-1798-3

Tables

Due to technical limitations, Table 1 is only available as a download in the supplementary files section.

Table 2. Baseline characteristics of the participants by sample			
Characteristic		Validity sample (n = 348)	Reliability sample (n = 133)
Age in years		66 ± 13 (18–90)	66 ± 14 (18–85)
Male/female		218/130 (63/37%)	75/58 (56/44%)
Disease-specific measures			
	Sub-acute stroke: NIH-SS, 0–42 points	5 ± 5 (0–22); n = 109	5 ± 5 (0–22); n = 51
	Parkinson's disease: Hoehn and Yahr scale, 0–5 stages	3 (3–4); n = 100	3 (3–4); n = 47
	Multiple Sclerosis: EDSS, 0–10 grades	6.5 (6–8); n = 18	6.5 (5–8); n = 8
Disease duration, ^a days			
	mean, SD (range)	30 ± 23 (7–150); n = 173	31 ± 23 (7–118); n = 63
	median (IQR)	22 (15–38); n = 173	22 (16–37); n = 63
Disease duration, ^a years			
	mean, SD (range)	11 ± 8 (1–44); n = 175	12 ± 19 (1–44); n = 70
	median (IQR)	9 (5–14); n = 175	10 (5–16); n = 70
Time since admission at baseline assessment in days			
	mean, SD (range)	7 ± 12 (0–87)	11 ± 15 (0–84)
	median (IQR)	4 (2–7)	6 (3–12)
Walking aid			
	None	162 (47%)	59 (44%)
	Rollator/walker	83 (24%)	37 (28%)
	One cane/stick	25 (7%)	12 (9%)
	Two crutches/walking sticks	14 (4%)	5 (4%)
	Not ambulatory/wheelchair	64 (18%)	20 (15%)
de Morton Mobility Index, points		58 ± 26 (0–100)	59 ± 25 (0–100)
NIH-SS: National Institutes of Health Stroke Scale; EDSS: Expanded Disability Status Scale; SD: standard deviation; IQR: interquartile range Values are mean ± SD (range), median (IQR) or absolute numbers (%) ^a Disease duration is reported in days for (sub-acute) conditions with an exact date of event, such as stroke or spinal cord injury. Disease duration is reported in years for chronic conditions such as Parkinson's disease or Multiple Sclerosis.			

Table 3. Summary Fit Statistics for Rasch analyses									
	Item residual		Person residual		Chi-Square fit statistic		Reliability		Unidimensionality
	Mean	SD	Mean	SD	Value and (df)	P	PSI	Alpha	t-test; % (95% CI) < 5%
Results	-0.47	0.47	-0.08	0.01	59.4 (45)	0.074	0.90	0.94	2.2% (0.2 - 4.7)
Ideal values	0	1.0	0	1.0		0.003 ^a	> 0.7	> 0.7	< 5.0%
^a : Bonferonni adjustment alpha level for the DEMMI scale with 15 items (0.05/15); SD = standard deviation; df = degrees of freedom; PSI = Person-item separation index; CI = confidence interval									

Figures

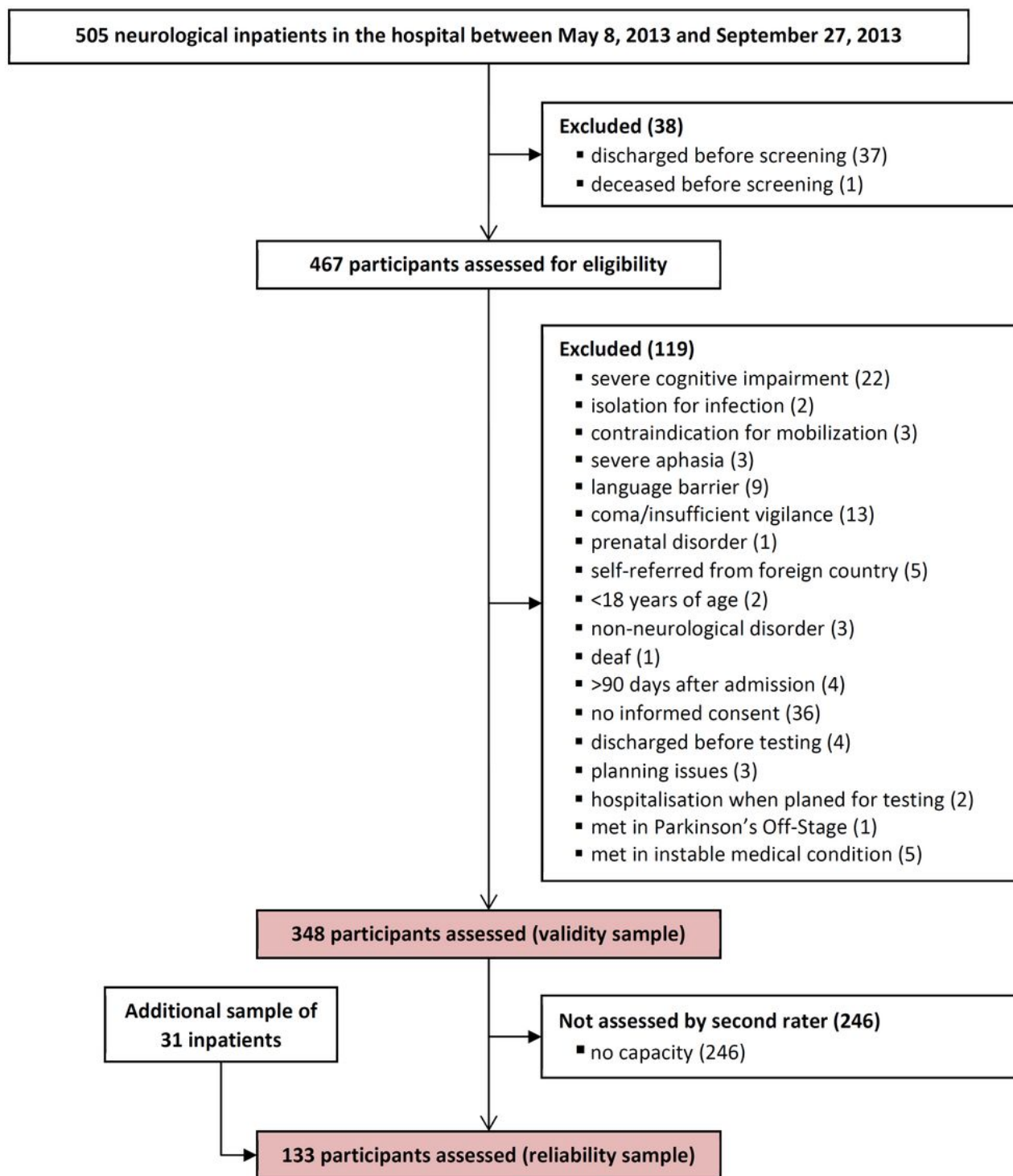


Figure 1

Figure 1. Flow chart of study participants.

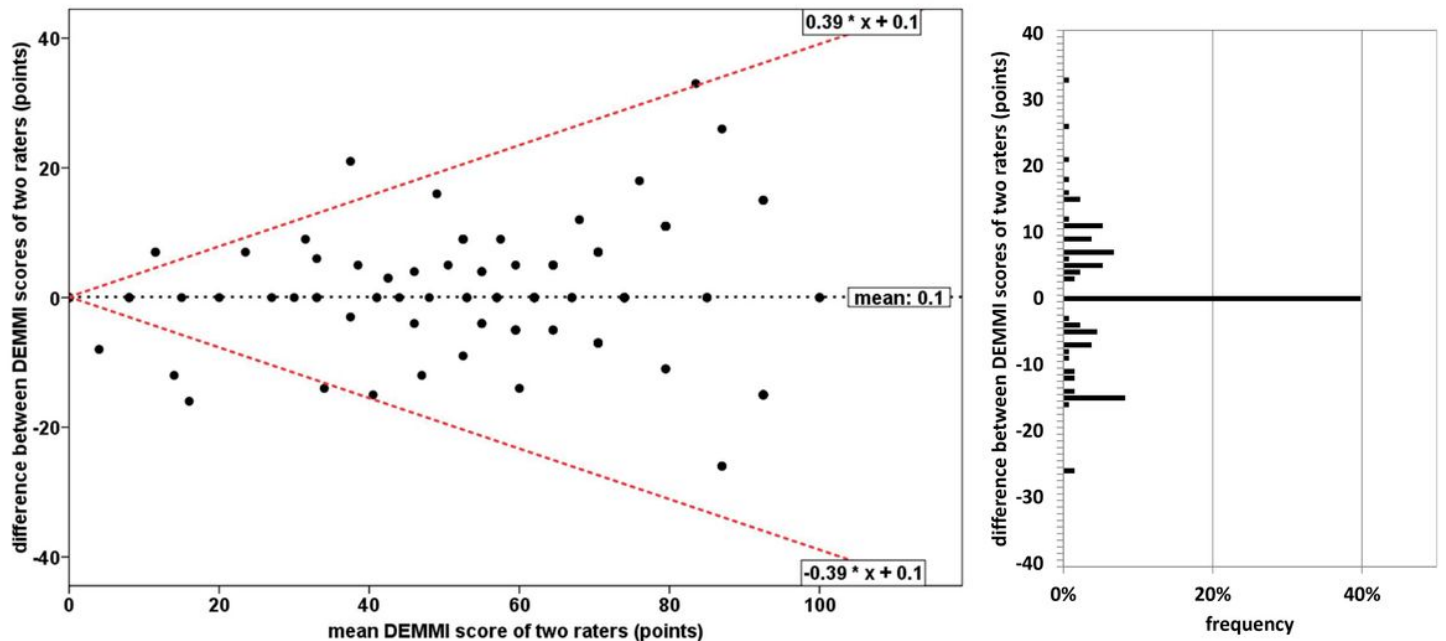


Figure 2

Figure 2. Bland-Altman plot of de Morton Mobility Index (DEMMI) scores by two raters. The x-axis represents the mean scores of the raters and the y-axis represents the difference between the raters. The dotted black line represents the mean difference between both measures; dotted red lines represent the 95% upper and lower limits of agreement. The bar chart on the right side illustrates the frequency of differences between the two raters.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Sup.file1Additionalmethods210317.pdf](#)
- [Sup.file2Additionalresults210317.pdf](#)
- [table1.docx](#)