DOI:
10.1080/09638288.2019.1643933

Document Version
Peer reviewed version

Link to publication record in King's Research Portal

Citation for published version (APA):
Ashford, S. A., Siegert, R. J., Williams, H., Nair, A., Orridge, S., & Turner-Stokes, L. (2019). Psychometric evaluation of the leg activity measure (LegA) for outcome measurement in people with brain injury and spasticity. *Disability & Rehabilitation*, 1-12. https://doi.org/10.1080/09638288.2019.1643933

Psychometric evaluation of the Leg Activity measure (LegA) for outcome measurement in people with brain injury and spasticity.

Stephen A Ashford[1, 2, 3], Richard J Siegert[4], Heather Williams[1], Ajoy Nair [5], Samantha Orridge[6], Lynne Turner-Stokes[1, 2]

[1]Regional Hyper-acute Rehabilitation Unit, Northwick Park Hospital, London North West Healthcare NHS Trust, London, United Kingdom.

[2]Department of Palliative Care, Policy and Rehabilitation, Faculty of Nursing, Midwifery and Palliative Care, King's College London, London, United Kingdom.

[3]Centre for Nursing and Midwifery led Research & Centre for NeuroRehabilitation, National Hospital for Neurology and Neurosurgery, University College London Hospital, London, United Kingdom.

[4]AUT University, Department of Psychology, Faculty of Health and Environmental Sciences, Auckland, New Zealand.

[5]Alderbourne Rehabilitation Unit, Hillingdon Hospital, London, United Kingdom.

[6]Kings College Hospital, London, United Kingdom.

Corresponding author:

Dr S Ashford, orcid.org/0000-0001-5541-7063, Regional Hyper-acute Rehabilitation Unit, Northwick Park Hospital, London North West Healthcare NHS Trust, Watford Road, Harrow, London, HA1 3UJ, United Kingdom. Stephen.ashford@nhs.net

Researchgate: StephenAshford

Linkedin: StephenAshford

Twitter: stephenashford3

www.researcherid.com/rid/D-6653-2013il

Declarations of interest

The first author, in conjunction with the co-authors undertook the initial development and construction of the LegA.

Funding

Stephen Ashford is funded by the National Institute of Health Research (NIHR) in the form of a Clinical Lectureship award (CAT-CL-03-2012-010) and NIHR CLAHRC Northwest London through a fellowship award. This paper presents independent research funded by the National Institute for Health Research (NIHR) UK and NIHR CLAHRC Northwest London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the NIHR CLAHRC Northwest London or the Department for Health, UK.

Authors' contributions

SA is senior author/chief investigator and conceived, designed, obtained funding for this work and wrote the drafts of the paper. RS made direct contributions to the psychometric analysis and writing. LTS made direct contributions to the written presentation of the paper and conception of the project. Heather Williams assisted with data handling, analysis and commented on the paper. AN assisted with data collection, recruitment and commented on the paper, he acted as principal investigator at one of the study sites. SO assisted with data collection, recruitment and commented on the paper, she acted as principal investigator at one of the study sites. All authors read and approved the final manuscript.

Availability of data and material

Data are available pertaining to this manuscript through the corresponding author from the Cicely Saunders Institute, King's College London data repository. The granted ethical approval and permissions do not allow for data to be stored in a fully open access repository.

Acknowledgements

We would like to thank patients, carers and clinical colleagues who helped with this work. In particular, thanks to Denise Hancock, senior physiotherapist for assistance with participant recruitment and support for measure completion.

Psychometric evaluation the Leg Activity measure

Abstract

Purpose

The aim of this study is to evaluate the properties of the Leg Activity measure according to COnsensus-based Standards for the selection of health Measurement INstruments.

Method

Participants were assessed at baseline, one day, 6 weeks and 12 weeks, following treatment for leg spasticity with botulinum toxin and physical interventions.

Results

In stage 1, 64 participants were recruited to evaluate the initial psychometric properties of Leg Activity measure. In stage 2, 100 additional participants were recruited, to evaluate the scaling properties. Total sample of 164 participants was used.

Construct validity was supported for ‘passive function’, ‘active function’ and ‘impact on quality of life’. Cronbach’s alpha was 0.86, 0.97 and 0.87 respectively for the scales. Item level test-retest agreement ranged from 91-97% (Kappa 0.75-0.95). Following treatment for spasticity (n=64), the Leg Activity measure ‘passive function’ and ‘impact on life’ scales were responsive to change.

Principal components analysis confirmed the constructs and a unidimensional Rasch Partial Credit Model was subsequently established for each scale. Transformation to an interval scale was achieved. Using the ordinal-to-interval conversion tables, parametric statistical analysis may be used.

Conclusion

The study provides support for the Leg Activity measure being valid, reliable and responsive.

Keywords

Leg, Psychometrics, Quality of Life, Stroke, Brain injuries

Introduction

Spasticity is a common and distressing consequence of acquired brain injury, including stroke, trauma and other forms of brain injury. It is estimated that up to one-third of post-stroke patients develop symptomatic spasticity [1, 2] (overall range 4-42% [2, 3]), and approximately 75% of patients with physical disability following severe traumatic brain injury will develop spasticity requiring specific treatment. Of these, approximately one-third may require treatment with botulinum toxin. Spasticity interferes with leg movement and limits active functional tasks such as mobility and transfers, as well as increasing the burden on caregivers assisting people with personal care otherwise referred to as passive function [4]. Goals for treatment of leg spasticity often focus on active function improvements in walking, standing and transferring from different seated positions. However, improvements in passive function tasks and symptom management are often just as important to patients. Active function is the use of the limb to directly perform a task. Passive function is the care of the affected limb, usually performed by the individual themselves, but may require assistance from another person [5].

There is a need for instruments with demonstrable measurement properties capable of reflecting clinically important change in practice. Outcome measures developed for treatments such as botulinum toxin and physical interventions for spasticity aimed at improving any elements of functional performance (active and passive function), should ideally reflect function in real life, as opposed to just observed tasks in the clinic setting. Patient-reported measures offer this advantage and minimise the burden of data collection for clinicians in the clinical environment.

Our systematic review of the published literature [6], identified a number of measures that addressed leg function (eight in total) using self-report by patients and carers. All the measures were designed for self-report by patients but some could also be completed by interview. All the identified measures addressed active function (within the activity domain of the International Classification of Function, Disability and Health [7]). However, none addressed passive function – i.e. caring for the limb. A new self-report measure of active and passive function in the leg was therefore developed, the Leg Activity measure [8]. Conceptualisation and development of the Leg Activity measure has been published previously [8], in this paper we present the evaluation of its psychometric measurement properties.

Outcome measures should be subject to evaluation confirming they provide a valid and reliable assessment of the clinical parameters in question and to allow understanding of their psychometric properties. The COnsensus-based Standards for the selection of health Measurement INstruments criteria provide a set of attributes against which the psychometric properties of health status and quality of life instruments may be judged [9-11]. We aimed to use this framework to provide an evaluation of the psychometric properties of the Leg Activity measure. Criteria addressed in this initial analysis were; internal consistency, reliability, structural validity, interpretability, construct validity (hypothesis testing), responsiveness, and feasibility.

Methods

At stage 1, initial psychometric evaluation was conducted in a cohort of participants, from three spasticity services. In stage 2, patients seen in routine practice for spasticity management were also recruited from one service.

Sample Size Calculation

The sample size was based on the criteria by Terwee and colleagues for evaluation of construct validity and test re-test reliability in groups of at least 50 participants [9] and recommendations by Sim and Wright for sample size in reliability studies [12]. This was applied to the recruitment of participants in stage 1.

In stage 2 larger numbers of participants were required to enable more robust analysis using Principal Components Analysis and Rasch analysis resulting in an expanded cohort for these analyses. According to Lundgren-Nilsson and Tennant (2011) sample size should be a minimum 20 cases per item in the largest subscale for Rasch analysis (16,26).

Stage 1: Subjects and setting

The sample for stage 1 psychometric analysis was a prospective consecutive cohort of patients with lower limb spasticity presenting for treatment at one of three specialist spasticity services in London, UK between November 2014 and June 2016. Inclusion in the cohort study was based on a clinically identified need for spasticity management (including botulinum toxin injection) as part of a rehabilitation programme. Services were chosen purposively to reflect the range/types of specialist spasticity services available in the UK, but with relevance more widely to other countries and health systems. The services were:

1. An academic regional rehabilitation service, with associated outreach (to other rehabilitation units, community and referring hospitals) and spasticity service.
2. A local specialist rehabilitation service with associated spasticity clinic.
3. A regional spasticity service based at a tertiary hospital.

The services represent the current range of services offering spasticity management in the context of rehabilitation in the UK.

The cohort was assessed at baseline (Time 1) and one day later (Time 2), for the evaluation of repeatability, and then at 6 weeks (Time 3) and 12 weeks (Time 4), following treatment for leg spasticity with Botulinum toxin and physical interventions.

Patients included were adults aged between 18 and 85 years with primarily a hemiplegic presentation and leg spasticity impacting either active or passive function. Patients were excluded if they declined to participate or if their family and/or treating team declined on their behalf; or if they were unable to complete a questionnaire and no carer (professional or family) was available to undertake questionnaire completion on their behalf.

Stage 2: Subjects and setting

The additional sample for stage 2 analysis was a prospective consecutive, routinely collected, cohort of patients with lower limb spasticity presenting for treatment at one specialist spasticity service in London, UK between June 2016 and March 2018. Inclusion in the cohort study was based on a clinically identified need for spasticity management (including botulinum toxin injection) as part of a rehabilitation programme.

The cohort was assessed at baseline (Time 1). Patients included were adults aged between 18 and 85 years with primarily a hemiplegic presentation and leg spasticity impacting either active or passive function. All participants underwent treatment for spasticity requiring Botulinum toxin intervention and physical interventions. Patients were excluded if they declined to participate or if their family and/or treating team declined on their behalf; or if they were unable to complete a questionnaire and no carer (professional or family) was available to undertake questionnaire completion on their behalf.

Measures

Measures in this study were used to support psychometric evaluation of Leg Activity measure. The Leg Activity measure, Rivermead Mobility Index and Euro Quality of life 5 Dimensions were used to rate function by patients on the basis of activities performed over the preceding seven days. The Modified Ashworth Scale is used to rate spasticity and Goal Attainment Scaling is used to measure goal outcome.

- The **Leg Activity measure** [8] is a patient or carer-rated 33-item measure of difficulty in passive and active leg function, as well as spasticity related quality-of-life (symptoms and participatory impact). It comprises a nine-item Passive function subscale, a fifteen-item Active function subscale and a nine-item symptoms and impact on quality-of-life subscale, and uses a Likert scoring system between 0 (No difficulty) and 4 (unable to do task). The Passive function scale scores range from 0 (high function) to 36 (no function), Active function scale scores from 0 (high function) to 60 (no function) and Symptoms and Impact scale from 0 (high function) to 36 (no function). The Leg Activity measure is freely available to use and can be obtained with full instructions for completion, from the King's College London, Department of Palliative Care, Policy and Rehabilitation web site: https://www.kcl.ac.uk/lsm/research/divisions/cicelysaunders/attachments/LEG ACTIVITY MEASURE.pdf .

- The **Rivermead Mobility Index** [13-16] is a fifteen item measure of mobility (active) function items in a hierarchical order with increasing item difficulty. Items are dichotomous with a yes / no response option. The Rivermead Mobility Index was used as a comparison measure with Leg Activity measure for the active function scale.

- The **Euro Quality of life 5 Dimensions [17]** is a five item questionnaire with a further vertical scale on general health, scored from 0 (worst health) to 100 (best health). The five items are each scored on a five-point scale from 'no health limitation' to the 'most extreme health limitation'. The Euro Quality of life 5 Dimensions was used as a comparison measure with Leg Activity measure for the impact on quality of life scale.

- The **feasibility questionnaire** was used to evaluate ease of use, relevance and value in the clinical situation for the Leg Activity measure. It comprises one question each for a) time to complete, b) relevance and c) ease of completion.

Each question is rated on a 5-point Likert scale (e.g. for ease of completion: Very easy, Easy, Moderate, Difficult, Very difficult).

- The **Modified Ashworth Scale** [18, 19] is a clinical measure of spasticity which forms a single item scale from 0 (no increase in muscle tone) to 4 (affected part rigid in flexion or extension), with an additional point at +1 (slight increase in muscle tone) producing a six-point scale. The Modified Ashworth Scale therefore provides a single score to represent spasticity. Modified Ashworth Scale was scored for the relevant joint or joints and used as a comparison measure with the Leg Activity measure scales.

- **Goal Attainment Scaling** [20] is a method of evaluating the extent to which patient's individual goals are achieved in the course of intervention. Scoring of Goal Attainment Scaling followed the approach proposed by Turner-Stokes [21], but was used in this study to identify 'responders' and 'non responders' to treatment for spasticity without calculation of the 'T' score. Patients were categorised as responders if they achieved a score of 0 to +2 for their primary treatment goal, and non-responders if they achieved a score of -1 to -2 [20, 21]. Goal Attainment Scaling was also used as a comparison with the different scales of Leg Activity measure when the primary goal category related to that scale.

Goals were set prior to intervention using the Goal Attainment Scaling method for negotiating and recording goals [20, 21]. The Leg Activity measure, Rivermead Mobility Index, Euro Quality of life 5 Dimensions, Modified Ashworth Scale and feasibility questionnaire were recorded at baseline. The Leg Activity measure was repeated one day later to enable the evaluation of test-retest reliability. Measures were then repeated at 6 weeks and 12 weeks to enable the testing of responsiveness following intervention.

Patients gave written informed consent (n=51). For those patients unable to consent, consultation and approval was given by the next of kin (consultee) and the treating team (n=33) in accordance with the ethical approval received. A cleaned, validated dataset was exported to SPSS v24, RUMM2030 and STATA v14 statistical packages for analysis.

Ethical approval

Ethical approval for the research programme was received; National Research Ethics Service Committee London - South East (14/LO/1340). Participants gave their consent to involvement in this study. No personal or individual data are presented.

Stage 1: Analysis

Floor and ceiling effects were assessed in the study population by considering the percentage of participants at either extreme of the subscales according to the criteria by Terwee and colleagues [9].

Cronbach's alpha (Time 1) was used to evaluate internal consistency. A positive rating for internal consistency was given when ratings for Cronbach's alpha were between 0.70 and 0.95 [9].

Reproducibility (test re-test reliability) of Leg Activity measure was evaluated between Time 1 (Baseline) and Time 2 (one day later) using Quadratic Weighted Kappa coefficients for test re-test reliability to comply with the recommendations of Terwee et al [9].

Construct validity was established by evaluating hypotheses predicted for the three scales of Leg Activity measure when compared with other measures using the Spearman rank order correlation coefficient for convergent and divergent validity.

- For the passive function scale it was hypothesised that change would be correlated with Goal Attainment Scaling ‘T’ score at 6 weeks. It was also hypothesised that Modified Ashworth Scale would be correlated with passive function at 6 weeks.
- For active function it was hypothesised that it would be correlated with the Rivermead Mobility Index at 6 and 12 weeks.
- For the impact scale it was hypothesised that Euro Quality of life 5 Dimensions would not be correlated due to its general quality-of-life focus and the very specific focus on spasticity related quality-of-life for the Leg Activity measure impact. However, it was hypothesised that it would be correlated with Modified Ashworth Scale at 6 and 12 weeks.

Responsiveness of the Leg Activity measure was evaluated between baseline, 8 and 16 weeks following Botulinum toxin-A injection. Responsiveness was determined by comparing the change in Leg Activity measure between responder and non-responder groups using the non-parametric Mann-Whitney U test. It was expected that the Leg Activity measure would identify a significant difference between the responder and the non-responder groups for passive function as defined by Modified Ashworth Scale and primary goal outcome (Goal Attainment Scaling) as a result of spasticity intervention. A significant change was not expected in the group for active function, due the smaller number of active function goals in this cohort, though some individuals were expected to make these changes [22, 23].

Stage 2: Analysis

The unidimensionality of the three scales of the Leg Activity measure was initially examined using exploratory factor analysis and then Rasch analysis for confirmation of the structure and its scaling properties.

Factor Analysis

For exploratory factor analysis we used principal components analysis with Varimax rotation as these typically provide clear and interpretable solutions. All factor analyses were completed with IBM SPSS v.24 software. The Kaiser-Myer Olkin test and Bartlett's Test of Sphericity were used to ensure that the correlation matrix was appropriate for factor analysis for each subscale. The decision as to how many components to rotate was based upon consideration of: (i) the number of components with eigenvalues > 1.0; inspection of the scree plot; and parallel analysis which compares the number of eigenvalues > 1.0 with the number expected by chance alone.

Rasch Analysis

For Rasch analysis we followed recommendations suggested by Lundgren-Nilsson and Tennant (2011) for improving the rigor of Rasch papers in the context of rehabilitation including the following:

1. Use of the Andrich Rating Scale versus the Partial Credit Model chosen according to the Likelihood-Ratio test. The Rating Scale model assumes the distances between thresholds to be the same across all items whereas the Partial Credit

allows for flexibility. A significant Likelihood-Ratio test indicates the data are suitable for the Partial Credit Model.

2. Examine several 'analytical pathways', including some with and without re-ordering disordered thresholds.
3. Creation of testlets or 'super items' to deal with local response dependency. Local dependency is identified by correlations in the residual (unexplained) variance among items. Locally dependent items are simply summed or combined to form a super-item.
4. Unidimensionality is tested using Rasch principal components analysis of the residuals and the equating test with paired *t*-tests across all participants.
5. Where present, Differential Item Functioning might require splitting the sample according to the relevant person factor (e.g. age, sex, diagnostic group, etc.).
6. Item removal only as a last resort (in order to maintain the clinical integrity of the instrument).
7. Where possible, production of a transformation table to convert raw scores to Rasch-transformed scores, thus encouraging clinicians to use interval scores.

We followed the above steps to deal with each of these issues, when they arose. We used the Likelihood-Ratio test, to determine whether the Rating Scale or Partial Credit Model for Rasch analysis was most appropriate. The summary statistics of the Rasch model were assessed based on the mean item and person location, individual item fit residual, the overall item-trait interaction chi-squared test/p value and the Person Separation Index (PSI), interpreted as follows:

- The mean item location is always set to zero.
- A mean person location (reflecting needs) of ±0.5 indicates a well-targeted scale.

- The overall mean values representing perfect fit for both item and person fit residual are 0.0 (SD=1).
- The item-trait interaction chi square reflects the overall fit of the data to the model's expectations and should be not significant (*p*>0.05).
- The PSI is a measure of internal consistency of the scale, similar to a Cronbach's alpha in classical test theory (30); PSI values above 0.7 are required for group use and above 0.8 for individual assessment.

We tested for item bias across important person factors such as age group (<65, ≥65), gender, and diagnostic category (i.e. stroke, traumatic brain injury, anoxic brain injury, brain tumour, multiple sclerosis, spinal cord injury). As it was desirable to keep the original structure of the three Leg Activity measure scales, item removal would only be considered as a last resort to improve the fit. The items at risk of deletion were those exhibiting significant misfit, i.e., excessive residual values (> ± 2.5) and a *p* value significant at the 0.05 level, with a Bonferroni adjustment for multiple tests[24]. Unidimensionality was tested using Rasch principal components analysis of the residuals and the equating *t* test. Unidimensionality of the scale is confirmed if significant *t*-test comparisons do not exceed 5% or if the lower bound of a binominal confidence interval computed for the number of significant t-tests overlaps the 5% cut-off point[25]. Finally, we followed the 10 quality indicators for evaluating the quality of reporting Rasch studies developed by the Rasch Special Interest Group of Outcome Measures in Rheumatology [26].

Interpretability was addressed through the converting of the interval 'logit' scale produced by Rasch analysis into the ordinal scale structure of the original scale. This allows interpretation of the ordinal scale structure in identifying true change as identified through the 'logit' or interval scale for this sample at any point on the scale.

Feasibility was evaluated (Time 1) using a self-completed questionnaire administered following Leg Activity measure completion. Patients and carers rated the timeliness, ease of use and relevance.

Results

In stage 1, results are presented for the initial 64 participants. In stage 2, the initial 64 are combined with the additional 100 participants, with a total number of 164.

Stage 1: results

A total of 65 participants consented to involvement in stage 1, one patient subsequently refused clinical intervention and was excluded, 64 participants were therefore included in the prospective cohort. Mean age was 51 (SD 17.4), 32 (50%) male. The demographic characteristics of participants are described in Table 1.

Table 1 Demographic characteristics of the study population stage 1 (n=64)

The medians and interquartile ranges for the measures used at baseline, 1 day and 6 weeks are shown in Table 2.

Table 2: Median and inter-quartile range (IQR) for the study measures

The Leg Activity measure passive function, active function and impact scores were distributed over the full range of each subscale; passive function (0 to 36), active function (0 to 60) and impact (0 to 36). In the active function scale, a complete range of scores was produced, however a ceiling effect occurred with 23% (n=15) of scores for active function at the maximum point on the scale (i.e. totally unable to perform all items). Completion of the three subscales of Leg Activity measure was achieved by 62 (97%) of participants in 15 minutes or under.

Internal consistency (time 1): Cronbach's alpha was 0.86 for the Leg Activity measure passive function scale, 0.97 for active function and 0.87 for impact scale, demonstrating high internal consistency in each of the scales.

Test re-test reliability: Quadratic weighted Kappa coefficients for the Leg Activity measure scale scores at time 1 and 2 were passive function 0.83 (Standard Error 0.076), active function 0.91 (Standard Error 0.076), and impact 0.82 (Standard Error 0.076).

Item-level agreement ranged from 92% to 98%, and Quadratic weighted Kappa co-efficients ranged from 0.68 to 0.95. The Kappa coefficients conformed to "substantial" or "almost perfect" criteria for all items [27].

Construct validity: As expected a significant correlation was seen between change in the Leg Activity measure 'passive function' and Goal Attainment Scaling 'T' score at 6 weeks (Rho -.350; p=0.013). A significant correlation was also identified as predicted between composite Modified Ashworth Scale and Leg Activity measure passive function at 6 weeks (Rho -.247; p=0.049).

In keeping with the hypothesis, there was a significant correlation between the Leg Activity measure active function and the Rivermead Mobility Index at 6 weeks (Rho=-.834; p=0.000) and 12 weeks (Rho=-.892; p=0.000) following intervention.

In the Leg Activity measure impact scale, no correlation was seen as hypothesised between the Euro Quality of life 5 Dimensions. Correlations were seen between change in the impact scale and change in Modified Ashworth Scale at 6 weeks (Rho -.266; p=0.046) and 12 weeks (Rho -.276; p=0.037) as anticipated.

Responsiveness: Significant differences were identified between responders and non-responders according to Modified Ashworth Scale in the passive function scale at 6 weeks (p = 0.05) and 12 weeks (p = 0.007). A significant difference was also identified between responders and non-responders according to Goal Attainment Scaling at 12 weeks (p = 0.049) for passive function.

For the active function scale a significant difference was identified between responders and non-responders according to Modified Ashworth Scale at 12 weeks (p = 0.002), but not at 6 weeks. A significant difference was not detected for the responder group according to goal, reflecting the smaller number of active function goals in this cohort.

A significant difference was also identified between responders and non-responders according to Modified Ashworth Scale by the impact on quality-of-life scale at 6 weeks (p = 0.047) and 12 weeks (p = 0.039).

The passive function and impact scales of Leg Activity measure demonstrated significant (Friedman's Test – Non-parametric ANOVA) change from baseline to 6 weeks and 12 weeks (outcome) following spasticity intervention at 0.05 significance. However significant changes were not seen for active function, which also corresponded with lack of change recorded by the Rivermead Mobility Index and reflected the smaller number of active function goals in this cohort.

Stage 2: results

For the 100 additional participants mean age was 49 (SD 18.7), 56 (56%) male. The demographic characteristics of additional participants are described in Table 3.

Table 3 Demographic characteristics of the study population stage 2 (n=100)

Principal Component Analyses

Leg Activity measure Passive Scale

A principal component analysis of the nine passive items showed evidence of two distinct factors with seven items loading very highly (i.e. > 0.80) on the first unrotated principal component (PC) and two items (Splint, Foot) with low loadings on the first unrotated principal component and high loadings (i.e. >.75) on the second. Varimax rotation produced a very similar pattern of loadings and this is illustrated in Figure 1 (Passive) which plots the loadings of the nine passive items on the two rotated factors. Both the ‘splint’ and ‘foot’ items have clear clinical differences from the other items which are more directly focused on ease of personal care.

Figure 1 Variamx rotation loadings for passive, active and impact scales

Leg Activity measure Active Scale

A principal component analysis of the 15 active function items suggested a strong general factor with the first component accounting for 75.5% of total variance and all 15 items loading > 0.60 on the first unrotated component. The first two components accounted for 84% of total variance and were the only two components with eigenvalues > 1.0. Rotation of two factors resulted in poor evidence for two factors with 9 out of 15 items loading > 0.40 on both factors.

Leg Activity measure Quality of Life Scale

A principal component analysis of the nine Quality of Life items of the Leg Activity measure revealed two components with eigenvalues > 1.0 that together accounted for 66% of the total variance in responses. Varimax rotation produced good evidence for two factors one related to symptoms (e.g. Pain, Spasms) and one to social life (e.g. Activities, Social, Work) as illustrated in Figure 1 (impact) which plots the loadings of the nine passive items on the two rotated factors. From a real-life perspective symptoms and social and work life make logical sense in being different factors, both of which relate to quality of life. These factors again, have strong face validity and make clinical sense.

Rasch analysis

Structural validity was examined using adherence to the Rasch Partial Credit Model, once the appropriateness of application of the model had been confirmed.

Leg Activity measure Passive Scale

Analysis combining both samples (n=164) indicated reliability of the Leg Activity measure Passive by person separation index (PSI) of 0.74. However, overall fit to the Rasch model was unsatisfactory with χ^2 = 241.96 (18), *p*<0.00 (Table 4). Items with disordered thresholds were identified for; hygiene, splint application, wheelchair, hoisting, underwear, bed positioning and footwear. These items were successfully rescored. Overall fit to the Rasch model was still unsatisfactory and items for splint and foot were removed from the scale. This mirrored findings for the principal component analysis with both the 'splint' and 'foot' items having clinical differences from the other items which are more directly focused on ease of personal care (see Figure 2). Overall fit to the Rasch model was then satisfactory with χ^2 = 25.45 (21), p<0.23 (Table 4 and Figure 3).

Figure 2 Leg Activity measure Passive Scale item threshold maps

Table 4 Leg Activity measure Passive function, Active function and Impact scale item fit and location (n=164)

The person-item threshold distribution demonstrated a clear distribution of persons across the range of the scale (see Figure 3). Similarly, a consistent distribution of items was also seen (see Figure 2).

Figure 3 Leg Activity measure person-item threshold plots

Leg Activity measure Active scale

Again, combining both samples (n=164) indicated reliability of the Leg Activity measure Active by person separation index (PSI) of 0.95. However, overall fit to the Rasch model was unsatisfactory with χ^2 = 97.66 (30), p<0.00 (Table 4). Items with disordered thresholds were identified for; transfer car, indoor walking, turning, stairs, obstacles, walking over carpet, walking outdoors, walking over rough ground and walking half a mile. These items were successfully rescored. However, fit to the model could not be achieved through re-scoring items or their removal. Six super-items were produced using the Principal Components Analysis for reference. The items appeared to form a single scale with increasing difficulty along a continuum of the items (see Fig 2, Active) , which also corresponded to the clinically based prediction. Items were therefore placed into 6 super-items based on the PCA and clinical interpretation, maintaining the hierarchy of the items, with 1. turning, lying, 2. sitting, transfer chair and transfer car, 3. stand, sit to stand, forming the first three super-items; and 4. walking indoors and turning, 5. stairs, obstacles, walking over carpet, 6. walking outdoors, walking over rough ground and walking over half a mile forming the final three super-items. These super-items make clinical sense with the resulting super-items representing increasing difficulty from turning in bed and lying in item 1, to outdoor mobility and walking greater distances in item 6. Overall fit to the Rasch model was then generally satisfactory with $\chi^2 = 49.2$ (36), p<0.07 (Table 4).

Leg Activity measure Impact on Quality of Life

Both samples combined (n=164), indicated reliability of the Leg Activity measure Impact scale by person separation index (PSI) of 0.85. However, overall fit to the Rasch model was unsatisfactory with $\chi^2 = 33.28$ (18)/0.01, $p<0.01$ (Table 4). Based on the Rasch findings and those from the principal components analysis, two super-items were produced. The first super item included; pain, comfort, caring, spasms and range of movement. The second super-item included; mobility, activities, social activities and work activities. These super items again made clinical sense with the first super-item representing symptoms relating to quality of life and the second to functional activities and wider participation. Overall fit to the Rasch model was then generally satisfactory with $\chi^2 = 4.38$ (4), p<0.36 (Table 4).

Structural validity: Preliminary comparison of Rasch logits to ordinal scales

The logits produced by the Rasch model were then converted back to the original ordinal scale structures as shown in Table 5 for Passive function, Table 6 for Active Function and Table 7 for Impact on quality-of-life.

Table 5 Conversion of raw Leg Activity measure scores to logits and then re-scored to the original scale (n=164) for the passive function scale.

Table 6 Conversion of raw Leg Activity measure scores to logits and then re-scored to the original scale (n=164) for the active function scale.

Table 7 Conversion of raw Leg Activity measure scores to logits and then re-scored to the original scale (n=164) for the impact (quality of life) scale.

When converted back to the original structure all points on the original ordinal scales do not reflect a change on the interval Rasch logit scale, which is clearly demonstrated in the conversion. Using the re-scored ordinal scale structure should allow a clinician to identify when real change has occurred on the original scale according to the interval level scale produced using the Rasch model.

Feasibility: Ease of completion was rated as ‘Very easy’ to ‘Moderately easy’ by 91% (n=58) of patients or carers in the initial sample of 64 participants. Completion of the Leg Activity measure was undertaken by 97% (n=62) of participants in 15 minutes or under. Relevance of the measure was rated by 91% (n=58) of respondents as ‘Very relevant’ to ‘Moderately relevant’.

Discussion

This initial evaluation of the psychometric properties of the Leg Activity measure in relation to the COnsensus-based Standards for the selection of health Measurement INstruments checklist and Medical Outcomes Trust Quality Criteria [9, 10], supports the measurement properties of the Leg Activity measure scales. Construct validity is supported, with confirmation of predicted correlations with comparison measures. Principal Components Analysis enabled exploration of the structure of each of the three scales. Items in each scale showed a relationship to each other, confirming that the scales again have acceptable construct validity for application in the clinical environment.

On inspection of individual item-fit and individual person-fit statistics, a good fit to the Rasch Model was found but required sub-test analysis for the active and impact on quality of life scales. Internal consistency and test-retest reliability, were excellent in the three Leg Activity measure scales. Responsiveness was identified in the passive function and impact on quality-of-life function scales. In the active function scale, responsiveness was also demonstrated, but would benefit from further evaluation. The Leg Activity measure scales were feasible for patients to complete in a timely manner. Table 8 provides a summary of the psychometric properties.

Table 8 Summary of Leg Activity measure psychometric properties.

In the stage 1 study group (n=64), changes in passive function and active function scales mirrored the changes seen for goals in these areas or constructs. In the full cohort used (n=164), fewer active function goals were however set and therefore active function changes were seen in a smaller number of individuals. The conversion of logits back to the original format provides an indication for this population of real change at different points on the original scale.

The Leg Activity measure is the only published self-report measure in the current literature that addresses passive and active function as well as impact on quality of life, of the paretic lower limb in a comprehensive manner. The Leg Activity measure is designed for self-completion, making it potentially useful for patient and/or carer completion at a clinic or return by post following clinic visits with low clinician burden. The Leg Activity measure has also been used as the basis for structured interview undertaken at the clinic visit, as well as completed though telephone interview. The findings presented here in the context of treatment for lower limb spasticity, provide initial psychometric support for its further testing and use.

While no other measurement tool of this type is available for lower limb evaluation, a number of other tools have been developed and used in the context of upper limb spasticity [28]. The Arm Activity measure [29-31] was developed by our research team to address the need to measure improvements in arm passive and active function following spasticity treatment. The Leg Activity measure is the first tool to address these issues of patient reported functional outcome in the context of leg spasticity. The Leg Activity measure develops these concepts further than has currently been undertaken for arm spasticity in considering the impact of improvement in arm passive or active function on wider quality of life.

Correlations were not seen between the Leg Activity measure Impact scale and Euro Quality of life 5 Dimensions. As expected, based on previous studies of botulinum toxin intervention [32, 33], change in quality of life was not detected by Euro Quality of life 5 Dimensions following botulinum toxin intervention in the leg, however significant changes were identified following treatment, by the Leg Activity measure Impact scale in this cohort at outcome compared with baseline, indicating it forms a sensitive condition specific measure of spasticity related quality-of-life in the leg.

The authors recognise the following limitations to this study.

Firstly, it is important for a patient reported measure such as Leg Activity measure to capture both passive and active function, which form separate domains. However due to the functional impairment of those requiring intervention for passive function limitation, a floor effect was seen in active function for 23% of respondents. However in this context this does not indicate a limitation of the measurement scale, but rather further justification to support the application of both scales. However in the study group, there were fewer individuals with active function goals and correspondingly fewer individuals changing as measured by the active function scale.

Secondly, some items in the Leg Activity measure passive scale were identified with disordered thresholds. These items were rescored by collapsing the number of response options, which resulted in ordered thresholds for all items. Once the items with disordered thresholds had been rescored, item bias was not identified for age, gender, aetiology or diagnosis. The distribution of persons and items across the range of the scale show the ability of Leg Activity measure passive function to target the full range of function in this domain. Sub-test analysis was required to fit the active and impact of quality of life scales to the Rasch Partial Credit model. A good fit was then achieved.

Thirdly, a lower correlation was shown between Leg Activity measure and Modified Ashworth Scale. However, while we would expect there to be a relationship due to the intervention reducing spasticity (measured by Modified Ashworth Scale), the Leg Activity measure is measuring a different construct, that of activity (function), and so the relationship will not be perfect.

Fourthly, further evaluation of responsiveness to look at minimally important clinical difference was not possible in this study. With the Rasch converted scale, any change in the Leg Activity measure scales should represent a meaningful difference and represents the approach taken in the current work to addressing this issue. However, in the future it would be helpful to test this using the Rasch converted scale.

Despite these limitations, the study provides support for the Leg Activity measure as a valid, reliable and responsive tool for the evaluation of treatment in leg spasticity including condition specific quality-of-life. Ongoing evaluation, including additional exploration of Leg Activity measure scaling and measurement properties is underway.

Conclusions

This study demonstrated that the Leg Activity measure is a valid, reliable and responsive tool for the evaluation of treatment in leg spasticity, including a condition specific quality-of-life. While further testing is valuable, the Leg Activity measure can be used in practice and research for evaluation of outcome and measurement of clinical improvements in functional performance in individuals with post-stroke and acquired brain injury spasticity.

References

1. van Kuijk A, Hendricks H, Pasman J, Kremer B, Geurts A. Are neuroradiological or neurophysiological characteristics associated with upper-extremity hypertonia in severe ischaemia in supratentorial stroke? Journal of Rehabilitation Medicine. 2007;39(1):38-42.
2. Watkins CL. Prevalence of spasticity post stroke. Clin Rehabil. 2002;16(5):515-22.
3. Urban P, Wolf T, Uebele M, Marx J, Vogt T, Stoeter P, et al. Occurence and clinical predictors of spasticity after ischemic stroke. Stroke. 2010;41(9):2016-20.
4. Royal College of Physicians, British Society of Rehabilitation Medicine, Chartered Society of Physiotherapy, Association of Chartered Physiotherapist Interested in Neurology. Spasticity in adults: management using botulinum toxin - National Guidelines. London: Royal College of Physicians, Clinical Effectiveness and Evaluation Unit; 2009.
5. Sheean GL. Botulinum treatment of spasticity: Why is it difficult to show a functional benefit? Trauma and Rehabilitation. 2001:771-6.
6. Ashford S, Brown S, Turner-Stokes L. Systematic review of patient reported outcome measures (PROMS) for functional performance in the lower limb. J Rehabil Med. In press.
7. WHO. International Classification of Functioning, Disability and Health. Geneva: World Health Organisation; 2002.
8. Ashford S, Jackson D, Mahaffey P, Vanderstay R, Turner-Stokes L. Conceptualisation and development of the Leg Activity Measure (LegA) for patient and carer reported assessment of activity (function) in the paretic leg in people with acquired brain injury. Physiotherapy Reasearch International. 2016;DOI: 10.1002/pri.1660.
9. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker JH, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34-42.
10. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford P, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010;19:539-49.
11. Mokkink L B, Terwee C B, Patrick D L, Alonso J, Stratford P W, Knol D L, et al. COSMIN checklist manual. Amsterdam, The Netherlands: VU University Medical Center; 2012.
12. Sim J, Wright CC. The Kappa statistic in reliability studies: Use, interpretation and sample size requirements. Physical Therapy. 2005;85(3):257-68.
13. Collen FM, Wade DT, Robb GF, Bradshaw CM. The Rivermead Mobility Index: a further development of the Rivermead Motor Assessment. Int Disabil Studies. 1991;13(2):50-4. PubMed PMID: 1836787.
14. Hsieh CL, Hsueh IP, Mao HF. Validity and responsiveness of the rivermead mobility index in stroke patients. Scand J Rehabil Med. 2000;32(3):140-2. PubMed PMID: 11028799.
15. Roorda LD, Green JR, Houwink A, Bagley PJ, Smith J, Molenaar IW, et al. Item hierarchy-based analysis of the Rivermead Mobility Index resulted in improved interpretation and enabled faster scoring in patients undergoing rehabilitation after stroke. Arch Phys Med Rehabil. 2012;93(6):1091-6. PubMed PMID: 22464741.
16. Roorda LD, Green JR, Houwink A, Bagley PJ, Smith J, Molenaar IW, et al. The Rivermead Mobility Index allows valid comparisons between subgroups of patients

undergoing rehabilitation after stroke who differ with respect to age, sex, or side of lesion. Arch Phys Med Rehabil. 2012;93(6):1086-90. PubMed PMID: 22464740.
17. The EuroQol Group. EuroQol-a new facility for the measurement of health-related quality of life. Health Policy 1990;16(3):199-208.
18. Brashear A, Zafonte R, Corcoran M, Galvez-Jimenez N, Gracies JM, Gordon MF, et al. Inter- and intrarater reliability of the Ashworth Scale and the Disability Assessment Scale in patients with upper-limb post-stroke spasticity. Archives of Physical Medicine & Rehabilitation. 2002;83(10):1349-54.
19. Wade DT. Measurement in neurological rehabilitation. Oxford: Oxford University Press; 1992.
20. Ashford S, Turner-Stokes L. Goal attainment for spasticity management using botulinum toxin. Physio Res Int. 2006;11(1):24-34.
21. Turner-Stokes L. Goal attainment scaling (GAS) a practical guide. Clinical Rehabilitation. 2009;23(4):362-70.
22. McCrory P, Turner-Stokes L, Baguley I, De Graaff S, Katrak P, Sandanam J, et al. Botulinum toxin A for treatment of upper limb spasticity following stroke: A multi-centre randomized placebo-controlled study of the effects on quality of life and other person-centred outcomes. Journal of Rehabilitation Medicine. 2009;41:536-44.
23. Shaw L, Rodgers H, Price C, Van Wijck F, Shackley P, Steen N, et al. BoTULS: a multicentre randomised-controlled trial to evaluate the clinical effectiveness and cost-effectiveness of treating upper limb spasticity due to stroke with botulinum toxin A. Health Technologies Assessment. 2010;14(26):1-113, iii-iv.
24. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. Brit Med J. 1995:310:170.
25. Smith EVJ. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. J Appl Meas. 2002;3:205–31.
26. Leung YY, Png ME, Conaghan P, Tennant A. A systematic literature review on the application of Rasch analysis in musculoskeletal disease -- a special interest group report of OMERACT 11. J Rheumatol. 2014;41(1):159-64. Epub 2013/10/17. doi: 10.3899/jrheum.130814. PubMed PMID: 24128781.
27. Landis J, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159-74.
28. Bhakta BB, Cozens JA, Chamberlain MA, Bamford JM. Impact of botulinum toxin type A on disability and carer burden due to arm spasticity after stroke: a randomised double-blind placebo-controlled trial. Journal of Neurology, Neurosurgery and Psychiatry. 2000;69(2):217-21.
29. Ashford S, Slade M, Nair A, Turner-Stokes L. Arm Activity measure (ArmA) application for recording functional gain following focal spasticity treatment. International Journal of Therapy and Rehabilitation. 2014;21(1):10-7.
30. Ashford S, Slade M, Turner-Stokes L. Conceptualisation and development of the Arm Activity measure (ArmA) for assessment of activity in the hemiparetic arm. Disabil Rehabil. 2013;35(18):1513-8. doi: 10.3109/09638288.2012.743602.
31. Ashford S, Slade M, Turner-Stokes L. Initial psychometric evaluation of the Arm Activity Measure (ArmA): a measure of activity in the hemiparetic arm. Clin Rehabil. 2013;27(8):728-40. doi: 10.1177/0269215512474942.
32. McCrory P, Turner-Stokes L, Baguley I, De Graaff S, Katrak P, Sandanam J, et al. Botulinum toxin A for treatment of upper limb spasticity following stroke: A multi-centre randomized placebo-controlled study of the effects on quality of life and other person-centered outcomes. Journal of Rehabilitation Medicine. 2009;41:536-44.

33. Turner-Stokes L, Baguley I, De Graaff S, Katrak P, Davies L, McCrory P, et al. Goal attainment scaling in the evaluation of treatment of upper limb spasticity with botulinum toxin: A secondary analysis from a double-blind placebo-controlled randomised clinical trial. Journal of Rehabilitation Medicine. 2010;42:81-9.

Table 1 Demographic characteristics of the study population (n=64)

Mean age (years)	51 (SD=17.4; Range 18-84)
Male/female ratio	32:32
DIAGNOSIS	
Stroke	34 (53 %)
Traumatic brain injury	10 (16 %)
Anoxic brain injury	5 (8 %)
Brain tumour	6 (9 %)
Multiple Sclerosis	8 (13 %)
Spinal cord injury	1 (1 %)

Table 2: Median and inter-quartile range for the study measures

Measure	**Time 1** **Baseline** Median (IQR)	**Time 2** **1 day** Median (IQR)	**Time 3** **6 weeks** Median (IQR)
Leg Activity measure Passive	10.5 (5-15)	10.5 (5-15)	5 (2-11)
Leg Activity measure Active	50 (25.5-59)	51.5 (26.2-59)	51 (24-59)
Leg Activity measure Impact	21 (13.5-25.5)	21 (15.2-27)	13.5 (7.2-19.7)
Rivermead Mobility Index	1 (0-7.7)	-	2.5 (2-7)
EQ-5D (Health scale 0-100)	60 (43.7-80)	-	65.9 (40-80)
Modified Ashworth Scale (Total)	13 (8-18)	-	7 (4-12)
Goal Attainment Scaling	37.6 (31.4-37.6)	-	50 (42.5-50)

Key: IQR - inter-quartile range

Table 3 Demographic characteristics of the study population (n=100)

Mean age (years)	49 (SD=18.7; Range 17-81)
Male/female ratio	56:44
DIAGNOSIS	
Stroke	49 (49 %)
Traumatic brain injury	34 (34 %)
Anoxic brain injury	4 (4 %)
Brain tumour/Infection	5 (5 %)
Multiple Sclerosis	8 (8 %)
Spinal cord injury	0 (0 %)

Table 4 Leg Activity measure Passive function, Active function and Impact scale item fit and location (n=164)

Scale/Analysis	**Item**		**Person**		**Item-Trait Interaction**	**PSI**	**Unidimensionality**
	Location Mean (SD)	Fit residual Mean	Location mean	Fit residual Mean	χ^2 (DF)/p-value		(% Signif. t-test)
Passive (9 items) Analysis 1	0.00 (0.32)	-0.46 (2.53)	-1.09 (1.11)	-0.23 (1.23)	241.96 (18)/0.00	0.74	Yes 0%
Passive (9 items rescored*) Analysis 2	0.00 (0.40)	-0.36 (2.61)	-1.48 (1.46)	-0.24 (1.26)	190.65 (18)/0.00	0.76	Yes 3.66%
Passive (7 items) Analysis 3	0.00 (0.55)	0.12 (0.82)	-1.74 (1.93)	-0.51 (1.54)	25.45 (21)/0.23	0.83	Yes 7.32% (CI 4 – 10%)
Active (15 items) Analysis 1	0.00 (1.14)	-0.27 (2.06)	1.80 (2.85)	-0.38 (1.21)	97.66 (30) 0.00	0.95	No 32%
Active (6 super items) Analysis 2	0.00 (1.30)	-0.46 (1.83)	1.32 (2.10)	-0.26 (0.82)	49.2 (36)/0.07	0.93	Yes 3.85%
Impact (9 items)	0.00 (0.50)	0.54 (1.89)	0.06 (1.11)	-0.46 (1.64)	33.28 (18)/0.01	0.85	No 13%
Impact (2 super items)	0.00 (0.23)	0.30 (0.45)	0.00 (0.57)	-0.44 (0.87)	4.38 (4)/ 0.36	0.85	Yes 1.5%

*Rescored item categories are indicated in figure 3; 7 item passive function threshold map

Table 5 Conversion of raw Leg Activity measure scores to logits and then re-scored to the original scale (n=164) for the passive function scale.

Ordinal Measure Leg Activity measure Raw score	Interval Measure	
	Logits (Rasch Location)	Re-score to 36
0	-3.304	0
1	-2.647	3
2	-2.224	4
3	-1.95	6
4	-1.746	6
5	-1.58	7
6	-1.44	8
7	-1.316	8
8	-1.204	9
9	-1.101	9
10	-1.004	10
11	-0.912	10
12	-0.824	10
13	-0.738	11
14	-0.656	11
15	-0.574	11
16	-0.494	12
17	-0.413	12
18	-0.332	12
19	-0.25	13
20	-0.164	13
21	-0.072	13
22	0.026	14
23	0.134	14
24	0.257	15
25	0.397	15
26	0.559	16
27	0.748	17
28	0.966	18
29	1.221	19
30	1.519	20
31	1.873	21
32	2.294	23
33	2.799	25
34	3.428	28
35	4.27	31
36	5.399	36

Table 6 Conversion of raw Leg Activity measure scores to logits and then re-scored to the original scale (n=164) for the active function scale.

Ordinal Measure Leg Activity measure Raw score	Interval Measure	
	Logits (Rasch Location)	Re-score to 49
0	-3.515	0
1	-3.515	1
2	-2.805	5
3	-2.6	7
4	-2.431	7
5	-2.279	8
6	-1.836	11
7	-1.836	11
8	-1.836	11
9	-1.68	12
10	-1.519	13
11	-1.354	13
12	-1.19	14
13	-1.031	15
14	-0.881	16
15	-0.741	17
16	-0.611	18
17	-0.49	18
18	-0.49	18
19	-0.276	19
20	-0.179	20
21	-0.088	21
22	-0.001	21
23	0.082	21
24	0.163	22
25	0.242	22
26	0.321	23
27	0.4	23
28	0.48	24
29	0.48	24
30	0.65	25
31	0.742	25
32	0.742	25
33	0.742	25
34	1.045	27
35	1.153	27
36	1.264	28
37	1.376	29
38	1.492	29

39	1.612	30
40	1.739	31
41	1.873	31
42	2.015	32
43	2.163	33
44	2.32	34
45	2.32	34
46	2.69	36
47	2.945	37
48	3.347	40
49	3.979	43

The score range could only be converted back to the range represented in these data, which was up to 49 (rather than the original 60). This results from the ceiling effect for the active function scale (see person-item distribution figure 3).

Table 7 Conversion of raw Leg Activity measure scores to logits and then re-scored to the original scale (n=164) for the impact (quality of life) scale.

Ordinal Measure Leg Activity measure Raw score	Interval Measure	
	Logits (Rasch Location)	Re-score to 36
0	-2.768	0
1	-1.488	8
2	-0.973	11
3	-0.766	12
4	-0.668	12
5	-0.603	13
6	-0.551	13
7	-0.508	13
8	-0.468	14
9	-0.432	14
10	-0.4	14
11	-0.367	14
12	-0.333	14
13	-0.302	15
14	-0.268	15
15	-0.235	15
16	-0.202	15
17	-0.166	15
18	-0.132	16
19	-0.096	16
20	-0.06	16
21	-0.022	16
22	0.017	17
23	0.056	17
24	0.098	17
25	0.144	17
26	0.194	18
27	0.251	18
28	0.318	18
29	0.399	19
30	0.507	19
31	0.664	20
32	0.913	22
33	1.288	24
34	1.785	27
35	2.437	31
36	3.291	36

Table 8 Summary of Leg Activity measure psychometric properties.

Attribute	Criteria	Evaluation
Validity	Face	Confirmed during pilot testing.
	Content	***Aim; population and target concepts:*** The Leg Activity measure is designed to provide a low burden measure of difficulty in active and passive function and impact on quality of life for patients undergoing spasticity management in the leg. ***Item selection and reduction;*** Item selection used a systematic review and patient selected items followed by Delphi consensus process with specialist clinicians and confirmed with a patient and carer advisory group. ***Interpretability of items:*** Understanding was confirmed during pilot testing [9].
	Criterion-related	No accepted gold standard measure for comparison currently exists for passive and active leg function. The impact on quality of life scale, was not correlated due to apparent lack of sensitivity to change of EQ-5D in this population.
	Construct (Correlation with other measures).	***Convergent***: The passive function scale correlated with change in spasticity and goal outcome, active function was correlated with the Rivermead Mobility Index (Rho -0.25) and Impact on quality-of-life was not correlated with EQ-5D, but was correlated with changes in spasticity all as predicted. ***Divergent:*** The passive function scale did not significantly correlate with the Rivermead Mobility Index or EQ-5D. The active function subscale was not correlated with EQ-5D. The impact scale was not correlated with the Rivermead Mobility Index.
Reproducibility	Agreement	Percentage agreement ranged between 92% to 98% for the three scales.
	Test re-test Reliability (Weighted Kappa > 0.70 [9])	Total Quadratic Weighted Kappa coefficients for the passive function scale was 0.83, active function scale was 0.91 and impact scale was 0.82.

Responsiveness		A significant difference was identified between responders and non-responders using the Leg Activity measure passive function, active function and the 'impact on quality of life' scales.
Interpretability		Following evaluation of conformity to the Rasch model a logit conversion of the original ordinal scale has been produced for the passive function scale.
Floor/ceiling effects		No significant floor or ceiling effects in the passive function or impact scales was identified, but a 23% ceiling effect in the active function scale in the patient group used in this evaluation was identified.
Feasibility and Burden	Time to administer	Time for completion of Leg Activity measure was 15 minutes or under in 97% of patients in this analysis.
	Ease of use	Ease of completion was rated as very easy, easy or moderate by 91 % of patients or carers.
	Relevance	Relevance of the overall scale was rated by 91% of respondents as very relevant to moderately relevant.
Alternative modes of administration		Leg Activity measure has been administered during testing as a self-completion questionnaire or as an interview (face-to-face or over the telephone).
Cultural and language adaptations		None currently available

4

Figures

Figure 1 Variamx rotation loadings for passive, active and impact scales

Passive **Active** **Impact**

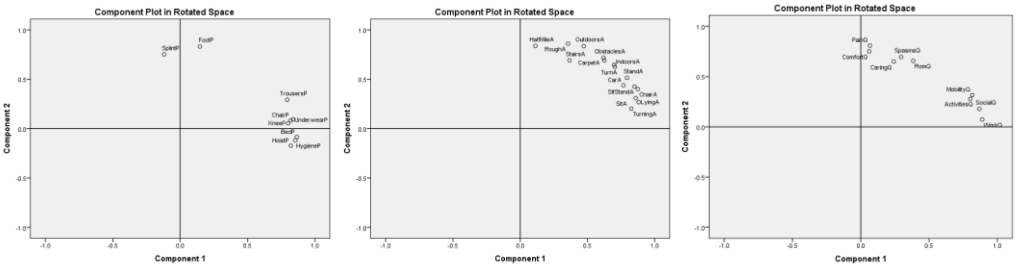
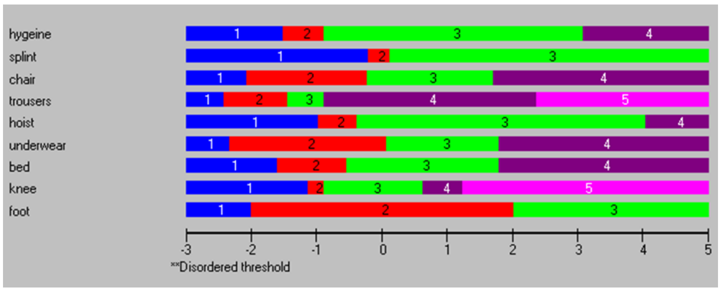


Figure 2 Leg Activity measure Passive Scale item threshold maps

Passive 9 Items



Passive 7 Items

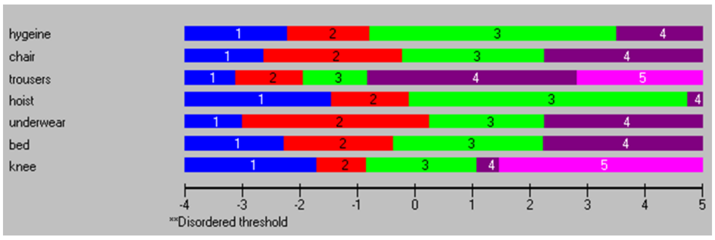
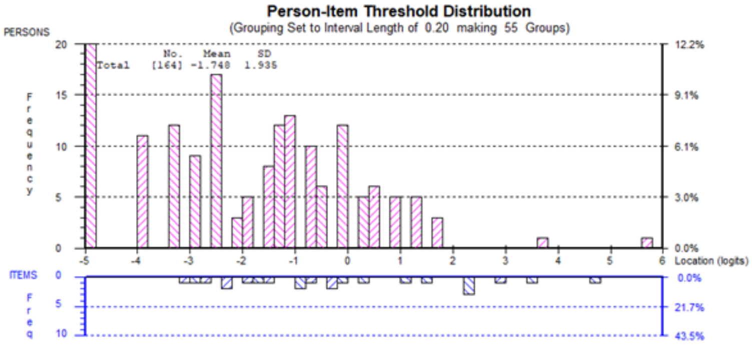
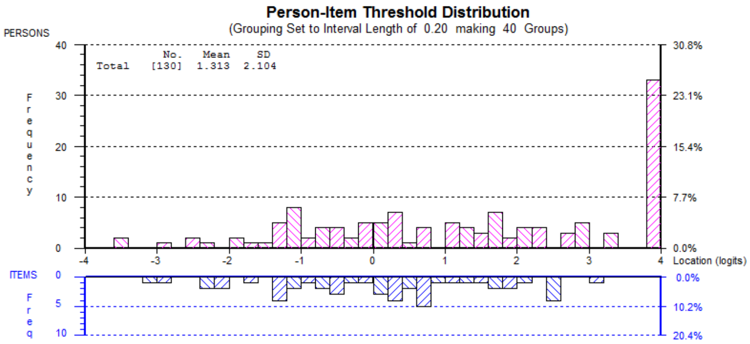


Figure 3 Leg Activity measure person-item threshold plots

Passive scale



Active scale



Impact on Quality of Life scale