Objective: To identify psychometrically evaluated patient-reported outcome measures reflective of ‘real-life’ function (active and passive) for application following focal rehabilitation interventions in the lower limb after stroke or brain injury.

Data sources: A literature search conducted in MEDLINE, CINAHL, Embase, Web of Science, PubMed, National Health Service National Research Register, MRC Clinical Trials directory, Database of Abstracts of Reviews of Effects (DARE), Google Scholar and the Cochrane Database of Systematic Reviews.

Study selection: Interventional study designs using patient-reported outcome measures for outcome evaluation meeting the review objective.

Data extraction: Evaluation of the psychometric properties used the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) process, by two independent reviewers reaching consensus, with adjudication by a third reviewer.

Data synthesis: One-hundred and thirteen studies were identified following initial review of the abstracts, yielding 12 outcome measures. Eight measures were identified, which were relevant to real life functional performance. These were the Brain Injury Community Rehabilitation Outcome, Climbing Stairs Questionnaire, Human Activity Profile, Lower Extremity Functional Scale, Nottingham Extended ADL Index, Rivermead Mobility Index (RMI), Sickness Impact Profile, Stroke Impact Scale.

Conclusions: All measures addressed active function, with none evaluating passive function. The RMI met most psychometric criteria, but may have a ceiling effect for high functioning patients.

Key words: systematic review; outcome assessment; psychometrics; lower limb; function; stroke; brain injuries.

INTRODUCTION

The rehabilitation process for patients who have suffered a neurological event such as stroke or brain injury should be focused on the needs of the individual and be person-centred (1, 2). The views of patients within this process of rehabilitation are particularly important (3, 4). Patient-Reported Outcome Measures (PROMs) are standardised, validated questionnaires that are completed by patients to measure their perceptions of their own functional status and wellbeing (5). PROMs have been identified as a method of involving patients in evaluating their ability and outcome following intervention (5, 6). PROMs also have the benefit of evaluating what patients do in their daily lives rather than only in the clinic setting and are therefore particularly relevant to the evaluation of person-centred outcomes (5, 7).

Brain injury or stroke often significantly impact motor function in the leg (8–11). Whilst many patients will recover to some degree of useful function in their lower limb, for a minority the limb effectively becomes a passive object to be cared for, either by the individual themselves or by a carer and may interfere with function (12–14).

Interventions for the lower limb may therefore be focused on a wide range of goals. At the higher level, interventions such as functional electrical stimulation (FES) or gait re-training may target recovery of mobility, walking and in some cases running. At a lower level, interventions such as spasticity management may be directed more towards goals in ‘passive’ function, such as making it easier to maintain perineal hygiene (14).

Outcome measurement is required to determine the effectiveness of rehabilitation interventions (15–17). Whether applied in clinical practice or for research, measures need to be valid, reliable and responsive to clinically relevant change (18, 19). Global measures of function in daily activities, such as the Barthel Index (14, 16), provide a general assessment of independence but are often unresponsive to focal interventions. Small changes which may be extremely important to the patient and/or their carers are easily lost amongst the larger number of unchanging items (20–22).

For these reasons, a number of motor function tests of gait, walking and balance have been developed, for example the
Berg balance test, 10-m walk test and the timed-up-and-go (TUG) test (23, 24). Conducted under close observation in the clinic, these may provide a more responsive and objective measure of motor activity (23, 24). However, they do not necessarily reflect how the person actually functions in terms of mobility in their normal environment, and it is generally not practical to obtain this information through 24-h observation in the home setting (2, 7). Instead, this information on ‘real life’ movement may be gathered through direct enquiry (self-report) from the patient and/or carer, for example using a task inventory administered by structured interview, self-completion questionnaire or PROM (25).

The aim of this systematic review was to identify valid and reliable patient (and/or carer) reported outcome measures that have been applied to assess changes following focal rehabilitation interventions in the lower limb in the context of stroke or brain injury, and are reflective of ‘real-life’ function (activity according to the International Classification of functioning, Disability and Health - ICF) for both active and passive function tasks (26).

Box: Active and passive function

**Active function:** Where a functional task is performed by active movement of the individual’s affected limb e.g. to stand, walk or actively participating in lower limb dressing.

**Passive function:** Where a task is carried out on the affected limb by the individual using the unaffected upper limb or by a carer e.g. cleaning the perineal area or putting on trousers or positioning the limb.

**METHOD**

The systematic review was performed by the authors in 3 stages, as described below, according to the methodology described by Ashford et al. (2). The review methodology is published and registered on the PROSPERO registry (CRD42013005046), Centre for Reviews and Dissemination, University of York. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement provides guidance on the most appropriate methods of presenting systematic review data and these principles were used, in the presentation of results (27).

Measure selection criteria were: 1) Application of the PROM in acquired brain injury (including stroke and traumatic brain injury). The stage 1 systematic search is used to identify measures applied in this area of practice. Studies are therefore not excluded on the basis of methodological design. 2) Measures are retained at stage 3 provided they address psychometric properties included in the COSMIN criteria (18). Psychometric evaluation is undertaken on all the retained measures. The exclusion criteria were: 1) Studies not evaluating any aspect of lower limb function. 2) Outcome measurement tools which are not PROMs. 3) Psychometric evaluation cannot be identified on the measures from the literature.

**Stage 1. Data sources**

In stage 1, a pool of possible measures was identified from a broad-based search of intervention studies. The following data sources were searched including Ovid MEDLINE, Ovid MEDLINE In-process, CINAHL, Embase, Web of Science, PubMed, National Health Service National Research Register, MRC Clinical Trials directory, Database of Abstracts of Reviews of Effects (DARE), Google scholar and the Cochrane Database of Systematic Reviews. The keywords used were leg, lower limb, hip, knee, ankle, foot, stroke, post stroke, cerebrovascular accident, CVA, cerebral haemorrhage, brain haemorrhage, haemorrhage, haematoma, hematoma, brain injuries, outcome measurement, outcome assessment, function (ing/al), activity, walking, transfers and mobility.

**Stage 2. Study selection**

The title was reviewed to identify potentially relevant studies. The abstract was then reviewed if the title was found to be relevant. When the abstract indicated relevance, the full text paper was retrieved and a final decision made about inclusion of the study. Initial selection was undertaken by the author and was then evaluated by a second reviewer; any areas of disagreement were discussed. Publications selected were restricted to those in the English language.

Measures were then excluded if they did not use a method of assessment reflective of ‘real-life’ function to measure day-to-day performance. In Stage 2, selected measures were considered to have ‘real-life’ relevance if they assessed day-to-day performance in the person’s normal environment, as opposed to performance when observed under test conditions (such as a standardised test in a clinic setting).

**Stage 3. Data extraction and synthesis**

In stage 3 a second systematic search was conducted, to enable evaluation of the published evidence for the psychometric properties of the selected measures, in addition to searching the reference lists of publications selected at stage 2. The COnsensus-based Standards for the selection of health status Measurement Instruments (COSMIN) checklist was applied in the study evaluation for assessing the methodological quality of the measurement instruments at stage 3 (18).

The names of measures identified in stage 2 were used as terms for a further search of the electronic databases to obtain original and any subsequent publications concerning their development and psychometric evaluation. MEDLINE, CINAHL and the reference lists of identified publications containing relevant outcome measures were then searched to identify further literature on the development of these outcome measures and their psychometric properties. Additional search terms used were: psychometric evaluation, testing, validity, reliability, application and clinical application. Authors of outcome measures were contacted for further details when required, in addition to searching the reference lists of the psychometric publications identified.

Based on this published literature, the psychometric properties of each measure were evaluated against the following review criteria based on the COSMIN process: Practicality for use in everyday practice: time to complete, burden, readability. Validity and reliability: content validity, internal consistency, construct validity, floor and ceiling effect, test-retest reliability, agreement. Responsiveness to change: demonstration of change following focal lower limb intervention, interpretability and minimal important change (MIC).

Descriptive information was tabulated for each of the selected measures including; the items in the measure, the methods of administration and the method of scoring applied. Two reviewers independently evaluated each measure using these criteria. Findings were then compared and any discrepancies resolved through discussion. The option was available for a third reviewer to resolve any areas of disagreement following comparison, but was not used.

**Procedure used to evaluate each measure**

The quality criteria developed by Bot et al. (19) were then used to operationalize the evaluation of the quality of each instruments properties, summarising each variable as adequate (+), doubtful (±), or poor quality (−), or as unknown (?) if insufficient information was available.

**Administrative burden**

Administrative burden was assessed using the same scoring method, modified as follows: Easy (+), when dichotomous items were simply summed; Moderate (+), when an ordinal or visual analogue scale was used to quantify individual items then summed, and Difficult (−) when a summary score was applied in combination with a formula. Timing
for completion of the measure was also rated as positive for measures completed within 10 min. Outcome measures should be both practical to use in routine practice and retain psychometric properties, thus ensuring the utility of the data produced (28, 29).

Validity
The instruments were evaluated for content and construct validity on the scale used for all psychometric properties. Content validity, evaluates that the instrument covers all the relevant concepts or domains (30). A positive rating for content validity was given when there was evidence that either patients, carers or other experts had been consulted regarding the initial selection of items (e.g. through focus groups or surveys) or had provided evaluation or feedback as part of the development. Construct validity originates from the idea that the new measure evaluates the construct it has been designed to measure (31). A positive rating for construct validity was given if there was evidence that the measure was based on hypothetical constructs, which had been tested and supported during its evaluation.

Internal consistency
Internal consistency refers to the interrelatedness of a set of items (32), and is often attributed to homogeneity of the items. A positive rating for internal consistency was given if the factor structure of the measure had been tested through factor analysis, or where ratings for Cronbach’s alpha were between 0.70 and 0.95 for each dimension or subscale.

Floor and ceiling effects
Floor and ceiling effects were considered present if more than 15% of respondents achieved the highest or lowest possible score, respectively.

Reliability (reliability)
Reliability is concerned with detecting the amount of error occurring during application of a measurement instrument (30). Test–retest reliability was rated as positive if repeat testing of the same condition had yielded comparable results, e.g. an intraclass correlation coefficient (ICC) of greater than 0.70 for total scores. In item-by-item analyses, agreement was also rated as positive if it had been evaluated and shown to be satisfactory, using accepted statistical methods such as the Kappa coefficient or standard error of measurement.

Responsiveness
Responsiveness is defined as ‘the ability of an instrument to measure a meaningful or clinically important change’ (33), when change occurs and record ‘no-change’ when the condition is stable. Responsiveness was rated as positive if the measure had demonstrated significant change in response to intervention, in the context of an appropriate study design.

Interpretability
Interpretability is the degree to which qualitative meaning can be assigned to quantitative scores (34). Positive ratings were given if at least two types of information were given to aid in understanding of the scores. Information considered included, means and standard deviations of the score totals before and after treatment, information in relation to other clinical variables, which might be expected to change, or information on the minimum change in score that might be clinically meaningful using the MIC (or MCID, Minimal Clinically Important Difference).

RESULTS

Stage 1. Data sources
A summary of the stages of review, according to the PRISMA measure selection flow diagram, is given in Fig. 1. The search yielded 19,942 studies, including primary reports, abstracts and conference proceedings. One hundred and thirteen studies were identified following initial review of the abstracts as including measures of functional outcome following focal lower limb intervention, yielding a total of 12 outcome measures after stage 1 (see Fig. 1).

Stage 2. Study selection
Eight measures were identified, which met both stage 1 and stage 2 criteria (i.e. were relevant to real life functional performance). These were the Brain Injury Community Rehabilitation Outcome (BICRO) (35); Climbing Stairs Questionnaire (36, 37), Human Activity Profile, Lower Extremity Functional Scale, Nottingham Extended ADL Index, Rivermead Mobility Index, Sickness Impact Profile, Stroke Impact Scale

Fig. 1. PRISMA Measure selection flow diagram.
Table I. Selected measures of function

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Method and procedure of scoring</th>
<th>Context for development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain Injury Community Rehabilitation Outcome Scales (BICRO)</td>
<td>- Items: 39 items; 8 domains (Personal care-6 items; Mobility-6 items; Self-organisation-6 items; Contact with partner/children-2 items; Contact with parents/siblings-3 items; Socialising-6 items; Productive employment-4 items; Psychological well-being-6 items) Scoring: 0–5 scale with variation in the descriptors according to the domain Administration: Patient self-report</td>
<td>Developed with patients with varied neurological conditions (TBI, Stroke, other ABI and MS)</td>
</tr>
<tr>
<td>Climbing Stairs Questionnaire (CSQ)</td>
<td>- Items: 15 items in one domain for ascending and descending stairs Scoring: Dichotomous variables Administration: Patient self-report or interview Modifications/versions: Original version in English only</td>
<td>Developed with patients with mixed lower limb impairment including stroke</td>
</tr>
<tr>
<td>Human Activity Profile (HAP)</td>
<td>- Items: 94 items in one domain of ‘activity’ Scoring: 0–3 scale Administration: Patient or carer self-report or interview Modifications/versions: Original version in English only</td>
<td>Developed with patients of mixed aetiology including stroke</td>
</tr>
<tr>
<td>Lower Extremity Functional Scale (LEFS)</td>
<td>- Items: 20 items in one domain of lower limb activity Scoring: 0–5 scale of task difficulty Administration: Patient or carer self-report or interview Modifications/versions: Original version in English and Portuguese</td>
<td>Developed in musculoskeletal problems, but applied in neurological conditions</td>
</tr>
<tr>
<td>Nottingham Extended ADL Index (N-ADL)</td>
<td>- Items: 21 items in 4 domains (mobility, kitchen, domestic, and leisure) relevant to activities of daily living (ADL) Scoring: 0–4 scale of undertaking task in past ‘few’ days Administration: Patient or carer self-report or interview Modifications/versions: Original version in English only</td>
<td>Developed with patients primarily with stroke (also multiple sclerosis)</td>
</tr>
<tr>
<td>Rivermead Mobility Index (RMI)</td>
<td>- Items: 15 items in one domain of lower limb activity Scoring: yes/no (0–1 scale) Administration: 14 self-report items and one clinician observation item (independent standing) Modifications/versions: Original version in English, Italian, Dutch, German and Portuguese. A modified version of the measure is available, but is entirely scored by the clinician and is not self-reported by the patient</td>
<td>Developed with patients with acquired brain injury</td>
</tr>
<tr>
<td>Sickness Impact Profile (SIP)</td>
<td>- Items: 30 items in 8 domains (Body Care and Movement, Social Interaction, Mobility, Communication, Emotional Behaviour, Household Management, Alertness Behaviour, Ambulation) Scoring: yes/no (0–1 scale) Administration: Patient self-report or interview Modifications/versions: Original 136 item version (SIP-136), which is burdensome to complete. A 68 item version (SIP-68) is also available. Original version and shortened versions in English and SIP-136 in German. An Italian version (SIP-23) for chronic pain is also available</td>
<td>Adapted from the SIP-68 to develop a new stroke specific version (SIP-30)</td>
</tr>
<tr>
<td>Stroke Impact Scale (SIS)</td>
<td>- Items: 60 items in 9 domains (physical problems, memory and thinking, mood, communication, activities, mobility, hand use, participate in activities, recovery) Scoring: 8 domains scored on a 5 point scale. 1 domain (recovery) scored with a single item on a numeric scale from 0 to 100 Administration: Patient self-report or interview Modifications/versions: SIS 2.0, SIS 3.0 and SIS-16. Translated into a number of different languages, originally developed in English</td>
<td>Developed with patients primarily with stroke</td>
</tr>
</tbody>
</table>

The administrative burden was adequate for 7 of the measures, but the HAP has a significant burden with time for completion (over 20 min in many studies) and contains 94 items. For the HAP construct validity was considered unclear according to the review criteria, with a resulting ‘doubtful’ rating.

DISCUSSION

In this systematic review of patient-reported measures of lower limb function, 8 measures were identified. All the measures had been applied in neurological populations, with 7 having undergone specific psychometric evaluation for this group (excluding the LEFS). Three of the measures specifically focused on evaluating lower limb function (CSQ, LEFS and RMI). The 5 remaining measures incorporated lower limb function in among self-report on other functional items or sub-scale dimensions. All the measures were self-reported by patients and some could also be completed by interview (see Table I). All measures identified following the search application, address active function within the activity domain of the ICF, with none evaluating passive function. Clinically-based measures of active function, such as the 10-m walk test or 6-min walk (24), are useful and valid.
PROMS for functional performance in lower limb measures of performance in the clinical environment under test conditions (105). However, they evaluate performance under an artificial test situation and not functional performance carried out under normal circumstances (real-life) by the individual (105). PROM’s are therefore very relevant in understanding active and passive function ability in the day-to-day environment, though clinically based measures maybe a complementary approach.

Of the 3 measures which focused specifically on lower limb function (CSQ, LEFS and RMI), the CSQ only evaluated stair climbing. The LEFS and RMI evaluated mobility in a number of different common situations and were broadly hierarchical in nature starting with easier items and progressing to more difficult items. The RMI was shown to be a unidimensional scale with a hierarchy of easy-to-hard items (65). The difficulty of items in the RMI was demonstrated to be stable when applied to different groups of patients assessed on different occasions in the same study. The RMI was developed and tested specifically with a population of patients with neurological impairment, while the LEFS was developed with patients with musculoskeletal problems, although it has been used in neurological populations. The RMI is therefore recommended as a hierarchical measure of lower limb active function, but may not address function at a very high level and therefore have a ceiling effect. The LEFS could be an option to evaluate high level function from a patient report perspective, but would need specific testing for patients with neurological impairment and therefore can not be recommended without further evaluation, which would be valuable.

Search strategy and PROM selection
The search strategy applied has enabled selection of measurement tools applied in research studies of patients with acquired brain injury including stroke. The use of two systematic searches has 1) facilitated the initial identification of relevant measures, and 2) been followed by a detailed evaluation of their published psychometric properties.

A range of patient-reported tools evaluating active function were identified, but no tool evaluating passive function. The search strategy focused on identifying patient reported tools, evaluating active and passive function. However, on reflection, the use of passive function as a search term, while appropriate, may have (in theory) excluded possible tools where authors

Table II. Psychometric evaluation from the literature of the selected measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Time</th>
<th>Admin burden</th>
<th>Content validity</th>
<th>Internal consistency</th>
<th>Construct validity</th>
<th>Floor/ceiling effect</th>
<th>Reliability</th>
<th>Agreement</th>
<th>Responsiveness</th>
<th>Interpretability</th>
<th>MCID</th>
</tr>
</thead>
<tbody>
<tr>
<td>BICRO</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>HAP</td>
<td>±</td>
<td>±</td>
<td>±</td>
<td>±</td>
<td>+</td>
<td>+</td>
<td>±</td>
<td>±</td>
<td>±</td>
<td>±</td>
<td>?</td>
</tr>
<tr>
<td>LEFS</td>
<td>+</td>
<td>+*</td>
<td>+*</td>
<td>+</td>
<td>+*</td>
<td>+</td>
<td>±</td>
<td>±</td>
<td>±*</td>
<td>±*</td>
<td>?</td>
</tr>
<tr>
<td>N-ADL</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>RMI</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>±</td>
<td>+</td>
<td>±</td>
<td>±</td>
<td>±</td>
</tr>
<tr>
<td>SIP</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>±</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>SIS</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>±</td>
<td>±</td>
<td>±</td>
<td>±</td>
<td>±</td>
<td>±</td>
</tr>
</tbody>
</table>

Method or result was rated as: + Adequate; ± Doubtful; – Poor; ? No data available.

*Not adequately evaluated in a neurodisability or acquired brain injury patient group.

**Demonstrated on lower limb sub-scales.

Admin: Administrative; BICRO: Brain Injury Community Rehabilitation Outcome Scales; CSQ: Climbing Stairs Questionnaire; HAP: Human Activity Profile; LEFS: Lower Extremity Functional Scale; MCID: Minimal Clinically Important Difference; N-ADL: Nottingham Extended ADL Index; RMI: Rivermead Mobility Index; SIP: Sickness Impact Profile; SIS: Stroke Impact Scale.

Table III. Categorisation of items included in each measure

<table>
<thead>
<tr>
<th>Item</th>
<th>BICRO</th>
<th>CSQ</th>
<th>HAP</th>
<th>LEFS</th>
<th>N-ADL</th>
<th>RMI</th>
<th>SIP</th>
<th>SIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turning in bed</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lying to sitting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sitting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transfer (bed to chair)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transfer (bath or car)</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sit to stand</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking indoors</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stairs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking outdoors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Running</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jumping/hopping</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endurance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

BICRO: Brain Injury Community Rehabilitation Outcome Scales; CSQ: Climbing Stairs Questionnaire; HAP: Human Activity Profile; LEFS: Lower Extremity Functional Scale; N-ADL: Nottingham Extended ADL Index; RMI: Rivermead Mobility Index; SIP: Sickness Impact Profile; SIS: Stroke Impact Scale.
had not used this term or concept. Other search terms such as ‘personal care’ could have been considered. We are, however, not aware of such tools from our research or clinical experience. The search did not include all possible search engines (e.g. PEDro), but did include indexes of publications most relevant to measurement in this area of outcome evaluation. Other measures may have been relevant to evaluate of active function. For example the ABILOCO (106), though applied via interview and not directly patient-reported, would nevertheless be relevant to assessment of this dimension. However ABILOCO was not identified from the stage 1 search or subsequent evaluation. Alongside considering an increase in the breadth of search terms applied, inclusion of other search engines could also be considered in future expanding on this review.

**Evaluation of psychometric properties**

The use of the COSMIN criteria in this systematic review has enabled a detailed and structured evaluation of the identified measures. The COSMIN criteria have been used as a tool to assess the measurement properties of various instruments. This systematic review used a combination of classical and item response methods to assess psychometric properties. In so doing, the COSMIN criteria make a useful contribution to the evaluation of patient-reported measures for clinically applied research and indeed practice.

In the current systematic review a wide range of psychometric evaluation of the different measures could be identified. In some cases, psychometric evaluation has been limited to one or two studies (e.g. BICRO); however for many measures evaluation was much more extensive. In some instances replication of psychometric evaluation had been possible largely confirming findings and supporting the properties being evaluated (e.g. RMI and N-ADL).

**Passive vs. active function**

While a number of measures addressing active function in the lower limb are available, no measure to evaluate passive function could be identified. This may be because active function improvement is more likely for greater numbers of patients in the lower limb, where smaller impairment improvement may lead to significant activity gain (e.g. small improvements in knee stability post stroke or brain injury may enable independent transfer and even walking).

Nevertheless from a clinical perspective, a group of patients can be identified who will not regain function and mobility. In a proportion of these patients, difficulty with issues such as spasticity and contracture may cause challenges in caring for the leg as seen in the upper limb (2, 7, 108). However, differences with the presentation in the upper limb are likely in the lower limb, and ease of care related to issues such as perineal hygiene may be more common in brain injury rather than stroke. For example, focal botulinum toxin intervention for lower limb spasticity (for example to hip adductors for perineal hygiene and catheter care), passive function improvement may be particularly important. In addition, passive function is possibly a more significant issue in the lower limb for people following spinal cord injury or those with severe multiple sclerosis. Passive function may be more significant in these groups, because of the bilateral nature of impairment in some cases and the severity of related symptoms such as spasticity. A review to consider if any additional PROMS are available in these patient populations would therefore be valuable.

**Summary and conclusions**

The RMI is a practical and clinically applicable measure of mobility in neurologically impaired, and in particular acquired brain injury, patients. The RMI has robust psychometric measurement properties, some of which have been replicated in a number of studies. Its dimensionality and measurement scaling properties have been evaluated and demonstrated using Rasch analysis and it can therefore be applied in a hierarchical manner to rate patient ability. It does however have a ceiling effect and high function patients may require a different patient-reported tool. The LEFS could be considered in high function groups, but has not been tested in neurological populations and is therefore not recommended at present.

In this systematic review, no measure was identified that addressed passive function. While a number of patient-reported measures of active function are available, the lack of passive function measures remains a deficiency. There is therefore a need for a measure that can evaluate passive function changes following intervention and management in the lower limb.

**ACKNOWLEDGEMENTS**

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

**REFERENCES**


60. Green S, Forester A, Young J. A test-retest reliability study of the Barthel Index, the Rivermead Mobility Index, the Nottingham extended Activities of Daily Living Scale and the Frenchay Activities Index in stroke patients. Disabil Rehabil 2001; 23: 670–676.


87. Lipsett PA, Swoboda SM, Campbell KA, Cornwell E, III, Dorman T, Pronovost PJ. Sickness Impact Profile Score versus a modified short-form survey for functional outcome assessment: acceptabil-


