How to measure pain in neurological conditions? A systematic review of psychometric properties and clinical utility of measurement tools

Sarah F Tyson and Philip Brown

Clin Rehabil published online 9 December 2013
DOI: 10.1177/0269215513514231

Article

CLINICAL REHABILITATION

Clinical Rehabilitation
0(0) 1–18

DOI: 10.1177/0269215513514231
cre.sagepub.com
SAGE

How to measure pain in neurological conditions? A systematic review of psychometric properties and clinical utility of measurement tools

Sarah F Tyson and Philip Brown

Abstract

Objective: To systematically review the psychometric properties and clinical utility of measures of pain in neurological conditions.

Data sources: Electronic databases (AMED, CINAHL, MEDLINE, PEDro and Web of Knowledge) were searched from their inception to February 2013.

Review methods: Studies investigating any measurement tool to assess pain in central nervous system conditions were systematically identified. Data about their psychometric properties and clinical utility were extracted and analysed independently. The strength of the psychometric properties and clinical utility were assessed.

Results: A total of 13 articles met the selection criteria, which assessed 11 measurement tools; eight pain rating scales; one Neuropathic Pain Scale; and two measures of pain interference with every-day life. Most of the pain rating scales were specifically for hemiplegic shoulder pain. None had been sufficiently developed to recommend for use in clinical practice or research. Evaluation of reliability and the ability to detect change were particularly sparse. Reliability depended on the type of tools used. Patients with right hemisphere damage favoured verbal/written responses, while people with left hemisphere damage preferred and reported more effectively using visual/numeric responses. Validity between measures of pain intensity was moderate, while validity with mood or quality of life was weak to moderate.

Conclusion: None of the selected measures of pain have been fully developed or evaluated to demonstrate that they provide accurate, relevant reproducible information.

Keywords

Pain assessment, neurological disorders, psychometry

Date received: 4 March 2013; accepted: 2 November 2013

Introduction

Pain in neurological conditions is thought to be common and important. Many patients identify it as one of their most troubling symptoms, but it is often described as incompletely understood, overlooked, and poorly managed.[1] For example, wide estimates of pain prevalence in different neurological conditions are found in the literature; between

Stroke and Vascular Research Centre, University of Manchester, Manchester, UK

Corresponding author:
Sarah Tyson, Stroke and Vascular Research Centre, School of Nursing, Midwifery and Social Work, University of Manchester, Jean McFarlane Building, Oxford Road, Manchester M13 9PL, UK.
Email: sarah.tyson@manchester.ac.uk

29% and 81% for people with multiple sclerosis;[2–4] 11%–94% for people with spinal cord injuries;[5] 41%–84% in adults with cerebral palsy;[6] 19% and 74% for stroke survivors;[7] 40%–85% in people with Parkinson's Disease;[8] and about 70% for people with Motor Neurone Disease.[9–11] This inconsistency is explained by the range of measurement tools used, patient populations recruited, and assessment protocols undertaken. At present, there is no widely accepted method to assess pain in people with neurological conditions, hence the drive for this systematic review, which aimed to assess the psychometric properties and the feasibility of use (clinical utility) so that recommendations regarding the most robust and easily used measures could be made.

Methods

The following electronic databases were searched; AMED, CINAHL, MEDLINE, PEDro and Web of Knowledge, from their inception to February 2013 using the following keywords:

- Pain
  AND
- assess* OR measure* OR tool OR outcome OR index OR test OR scale
  AND
- stroke *OR* cerebrovascular accident *OR* hemiplegia *OR* hemi* *OR* parkinson* *OR* sclerosis *OR* head injury *OR* brain injury *OR* guillian-barre *OR* motor neurone disease *OR* neuro*
- In the *PEDro* database we searched 'pain' AND 'neurology'.

Studies published in English were considered for inclusion if they assessed the psychometric properties of a tool to measure pain in adults with a central nervous system condition. The psychometric properties considered were:

- validity (concurrent or criterion related; construct or content);
- reliability (inter-rater or test–retest);
- ability to detect change (measurement error, standardized response mean, standardized error of measurement, limits of agreement, minimal detectable change, or ability to detect change during treatment).

The methods to assess these properties were as follows.

- For concurrent or criterion-related validity: (parametric or non-parametric) correlation coefficients.
- For construct or content validity (for ordinal scales):
  - internal consistency (Cronbach's alpha) or factor/ principal component analysis;
  - scaling properties: Rasch or Mokken analysis or co-efficients of reproducibility and coefficients of scalability (for hierarchy); inter-item correlations (for redundancy of items);
  - floor and ceiling effects.
- For reliability: intra-class correlations (for parametric data) and kappa statistics (for non-parametric data) or percentage agreement.
- For ability to detect change: measurement error, standardized response mean, standardized error of measurement, limits of agreement and minimal detectable change, or comparisons between groups or change over time.

Studies were excluded if they included measurement tools that were a composite of different constructs where the data for the pain-related items or sub-scales could not be extracted (measures of quality of life or condition severity, for example); measured psychometric properties or used methods other than those listed above; included a sample where less than 50% of the participants had a central nervous system condition or the data for the relevant participants could not be extracted; involved patients who did not have a central nervous system condition, such as peripheral neuropathy or generalized chronic pain.

To select the articles that met the inclusion criteria, the titles, then abstracts and full texts were independently screened by the authors. Consensus over selection was achieved through discussion, with a third party available to arbitrate if needed. As well as the databases, reference lists of the articles selected for the full text screening were searched and, finally, the databases above were searched using the names of the selected measurement tools. If necessary we contacted the original authors for

clarification regarding eligibility and for further data.

A description of the selected measurement tools, the participants, and data about the psychometric properties and clinical utility were independently extracted by the authors from the selected articles. If the results of the extraction or analysis differed, disagreements were resolved through discussion. A third person was available to arbitrate if agreement could not be reached.

First, the clinical utility of the measurement tools that met the selection criteria was assessed to quantify the practicalities of using the measurement tools. Previously developed criteria based on the factors that influenced whether clinicians would use a measurement tool in clinical practice[12] were used. These were as follows.

- Time taken to administer, analyse, and interpret the measurement tool: 3 < 10 minutes; 2 = 10–30 minutes; 1 = 30–60 minutes; 0 > 1 hour.
- Cost: 3 < £100; 2 = £100–500; 1 = £500–1000; 0 > £1000.
- Does the measurement tool need specialist equipment and training to use? 2 = no; 1 = yes, but simple and clinically feasible; 0 = yes and not clinically feasible/unknown.
- Is the measurement tool portable? Can it be taken to the patient? 2 = yes easily (can fit into pocket); 1 = yes (in a briefcase or trolley); 0 = no or very difficult.
- Is the measurement tool accessible? Can a detailed instruction for application be obtained? 2 = yes (full standardized operating procedure/instruction manual can be obtained from the article or a website); 1 = no, but operation can be simply worked out from a description in the article; 0 = no operating instructions available.

These scores were summated with a maximum of 12. Measurement tools that score less than 10 were considered infeasible for use in clinical practice and were rejected at this stage. The remaining measurement tools were considered feasible and their psychometric properties were assessed to identify those which would provide robust data.

The strength of the psychometric properties were assessed using a previously described framework[12–14] as follows.

+ weak reliability, validity, or internal consistency = scores of < 0.4
++ moderate reliability, validity, or internal consistency = scores of 0.4–0.6
+++ good reliability, validity, or internal consistency = scores 0.6–0.8
++++ excellent reliability, validity, or internal consistency = scores > 0.8

As data from the tests of ability to detect change, content, and construct validity and scaling properties are non-standardized, the acceptable (or unacceptable) limits were not specified but considered individually. The measurement properties were summarized to aid selection for use in practice. A measurement tool needed to obtain 'good' or 'excellent' scores for reliability and validity, and have some information about the ability to detect change before it could be recommended for use in clinical practice.

Results

The searches revealed 13 articles that met the selection criteria, which assessed 11 measurement tools (summarized in Figure 1 and Table 1). Eight were rating scales in which the patient identifies the intensity of their pain on a scale between no pain and the worst pain. The way in which the scale was presented varied. In a 'traditional' visual analogue scale (VAS), the scale is 10 cm long and the participant indicates where their pain lies on the scale and that point is measured, giving continuous ratio level data in millimetres.[15,17,19,20] Other designs convert the continuous scale into interval categories by adding numeric intervals (usually 0–10 or 0–5),[15,16,18] ordinal categorical data using descriptors (such as no pain; mild; moderate; severe or very severe pain),[15–17] or images (such as tick marks or faces with different expressions representing pain).[15–17,21,23,24] These can be presented vertically or horizontally, verbally or in writing; on plain paper or with a colour grading (with the colour becoming more intense/dark to represent more severe pain). The terms used to describe these

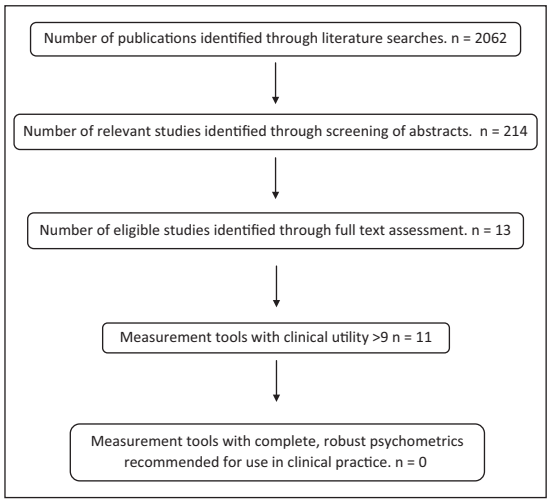

Figure 1. Showing the flow of articles through the systematic review.

different designs are inconsistent. The scales with descriptors have been called verbal rating scales or Likert scales (although it is not clear whether the tool is actually presented verbally or in writing); and the scales with numbers are referred to as numerical rating scales or visual graphic rating scales, and those with images as the Faces Pain Scale or Visual Grading Scales.

The severity or intensity of pain was the most frequent construct measured, but two scales sought to evaluate the impact of pain on patients' everyday lives, which was referred to as pain interference. These were the Pain Interference section of the Brief Pain Inventory and the Graded Chronic Pain Disability Score.[25–27] One tool specifically assessed neuropathic pain,[1] while the others focussed on nocioceptive/musculo-skeletal pain. Of these, hemiplegic shoulder pain was the most frequent target.[15,17,19,20,22–24] One measure was for hip pain in adults with cerebral palsy[21] and another article assessing several tools for unspecified pain.[16] None of the tools had been tested on a wide range of neurological conditions. The Neuropathic Pain Scale has only been tested in people with multiple sclerosis,[1] while the tools tested with survivors of stroke and acquired brain injury focused on shoulder pain.[15,17,19,20,22–24] The measures of pain interference had been most extensively tested, involving adults with cerebral palsy, spinal cord injury, and multiple sclerosis.[18,25,26] Pain in other common neurological conditions, such as Parkinson's disease, Motor Neurone Disease, head injury, or a generic neurological population is un-tested.

All the tools scored above 10 on the assessment of clinical utility. They were freely available, simple 'paper and pen' tests and scored full marks on the criteria for cost, portability, and need for specialist equipment. Time to complete the tests was not addressed in the articles, but from the details given one would expect them to be quick to perform. However, one measure of hip pain in people with cerebral palsy[21] involved a complex assessment of pain intensity during different activities, which would be time consuming when used for people with limited communication skills. Access to the tools' operating instructions was limited. Useable full operating instructions can be obtained for the multi-item ordinal scales from the authors,[23,24] other publications,[1] or the web,[18,25,26] but the pain rating scales[16–20] did not include operating instructions, although one could, to some extent, work out how the tools were intended to be used from their descriptions in the selected articles.

The psychometric evaluation is summarized in Tables 2 and 3, and the data extracted from the selected articles are detailed in Appendix 1. All the selected measurement tools had incomplete development and evaluation of the psychometric properties, particularly reliability and the ability to detect change (Table 2). Only one article[19] assessed all the reliability-related properties (using vertical VASs of physiotherapist-graded evaluation of pain) and found them inadequate as standard operating instructions were not used, nor had the assessors been trained how to use the tools/perform the assessment before testing. Where it had been tested, reports of reliability were varied but most indicated moderate levels of reliability. There was a difference between people with right- and left-sided brain damage.[17] Three measures had no published assessment of reliability in the selected neurological conditions; the Ritchie Articular Index,[20,22] Pain Interference Scale of the Brief Pain Inventory,[18,25,26] and Graded Chronic Pain Disability Score.[18,25,26]

Table 1. Measures of pain selected from the literature search with details of the type of measure and the clinical utility.

Measurement tool	Type of measure
Faces Pain Scale (6 and 7 level versions)[15–17]	Ordinal scale. Patients locate the severity of their pain on a vertical scale with 6 or 7 different facial expressions, expressing increasing levels of pain/distress. Different time scales are used: no time frame[15,17] or 'average pain over the 24 hours'.[16]
Likert Pain Scale or Verbal Rating Scales (VRS)[15–17,19,20]	Ordinal scale. Patients rate the severity of their pain against a series of descriptors (e.g. 'mild', 'moderate', or 'severe') where no pain is the lowest and the worst imaginable pain is the highest score. The choices may be presented verbally or in writing, or both. Different time scales are used: no time frame[15,17] or 'average pain over the 24 hours'.[16]
Numeric Rating Scales (NRS)[15–19]	Interval scale. Patients rate the severity of their pain against a series of numbers (0–10 or 0–100), where no pain is the lowest and the worst imaginable pain is the highest score. Different time scales are used: no time frame[15] or 'average pain over the 24 hours'.[16]
VAS (horizontal or vertical)[15,17,18,20,23,]	Continuous, ratio data. Patients rate the severity of their pain on a 10 cm line anchored at 0 (no pain) to 'worst possible pain' (10 cm), generally with no other indicators presented on the measurement line and no specified time scale.
Physiotherapist graded VAS (vertical)[19,22]	Continuous, ratio data. Physiotherapists graded the severity, frequency, and 'bothersomeness' and site of patients' hemiplegic shoulder pain following their clinical assessment based on patients self-report of pain; muscle activity, posture, presence of subluxation; emotional state and carers' report. No time scale specified.
Neuropathic Pain Scale[1]	Ordinal scale. Nine items scored on an 11-point (0–10) Numeric Rating Scale (horizontal with anchor points describing 'none' and 'worst imaginable' pain). Items asked about intensity and unpleasantness of the pain and how sharp, hot, dull, and cold the pain feels, plus how sensitive and itchy the skin feels and a descriptive item regarding the frequency of the pain. No time scale specified.
Pain Assessment Instrument for cerebral palsy (for hip pain)[18,21]	Ordinal scale measuring hip pain in adults with cerebral palsy. 12 diagrams of activities; six usually cause hip pain and six do not. Severity of pain in each activity is rated using the FPS-7. Patients are asked whether they have experienced pain during the activity and how much it hurt. No reference to time scale. The location of the pain also rated using a body map.
Ritchie Articular Index (for shoulder pain)[20,22]	Ordinal scale assessing (shoulder) joint tenderness. The assessor rates the patient's response to passive movement (external rotation and abduction) of the shoulder joint on a 4-point single-item scale: 0 = no pain; 1 = complains of pain; 2 = complains of pain and winces; 3 = complains of pain, winces, and withdraws.
ShoulderQ (for shoulder pain)[23,24]	Ordinal scale of shoulder pain intensity and timing using a mixture of formats. One 'yes/ no' question ('do you have shoulder pain?'). Two items scored with a 4-point Likert scale ('when do you have pain?' and 'how severe is your pain overall?'. Four items scored with a 3-point Likert scale ('does your pain wake you at night?', 'how many times per night?', 'does the pain interfere with therapy?', 'If so now much?'. One item scored on a 5-point Likert scale ('how severe is your pain compared with last week?'). Two 6-point Likert scales ('which tasks increase your pain?' and 'which interventions relieve your pain?'. Three items scored on a numeric rating scale (0–10) regarding severity of pain at rest, on movement, and at night
Pain Interference Scale of the Brief Pain Inventory[18,25,26]	Ordinal Scale. Items scored on a numeric rating scale (0–10) and then a composite (averaged) score of all the items is calculated. 7 items (the original tool) ask about the impact of pain on general activity, mood, mobility, work, relationships, sleep, enjoyment of life over the previous week. 10-item version added items regarding pain interference on self-care, recreational activities, social activities. 12-item version also asked about interference on cognition and communication.
Graded Chronic Pain Disability Score[18,25,26]	Ordinal Scale. Three items using a numeric rating scale (0–10) to score how the pain interfered with daily activity; recreational, social and family activities and the ability to work in the previous week'. Responses are averaged to produce a composite score.

FPS: Faces Pain Scale; VAS: visual analogue scale.

Table 2. Summary of the reliability of tools measuring pain in neurological conditions.

Scale	Assessed with	Construct measured	Test–retest reliability	Inter-rater reliability	Responsiveness/ sensitivity to change
Faces Pain Scale-6 or 7[15–17]	Adults with CP	Pain intensity	Not tested	Not tested	Not tested
	Stroke	Shoulder pain intensity	++ LHBD; ++ RHBD	++ LHBD; + RHBD	Not tested
NRS[15,16,18]	Adults with CP	Pain intensity	Not tested	Not tested	Not tested
Verbal Rating Scale/Likert Scale over 5 or 16 points[15–17]	Adults with CP	Pain intensity	Not tested	Not tested	Not tested
	Stroke	Shoulder pain intensity	+ LHBD; ++ RHBD	++ LHBD; ++RHBD	Not tested
VAS/NRS[15–17]	Stroke	Shoulder pain intensity	++ LHBD; + RHBD	+++ LHBD; +++ RHBD	Not tested
Horizontal or vertical VAS[15,17,19,20]	Stroke	Shoulder pain intensity	Intensity ++; frequency ++/+++; bothersomeness ++	Intensity ++/+++; frequency ++/+++; bothersomeness ++	Intensity ±60 points; frequency ±80 points; bothersomeness ±77 points
Pain Assessment Instrument of CP[21]	Adults with CP	Hip pain intensity	++	Not tested	Not tested
ShoulderQ[23,24]	Stroke	Shoulder pain severity	++	Not tested	Differences between responders and non-responders to treatment. Changes in total score seen during rehabilitation. Change >3 has 93% PPV to identity a responder to treatment
Neuropathic Pain Scale[1]	Multiple sclerosis	Neuropathic pain severity	Total and individual items ++	Not tested	Limits of agreement ±14 points

CP: cerebral palsy; LHBD: left hemisphere brain damage; NRS: Numerical Rating Scales; PPV: positive predictive value; RHBD: right hemisphere brain damage; VAS: visual analogue scale.

The validity of the tools had received more attention (Table 3). Where assessed, the tools were uni-dimensional and, where factor analysis had been used, demonstrated that pain intensity and pain interference were separate constructs,[15–17,25–27] while neuropathic pain was based on three constructs; alien, familiar, and superficial pain perceptions.[1] None of the ordinal scales had undergone evaluation against item-response theory. Inter-item correlations demonstrated that there was no redundancy of items in the Neuropathic Pain Scale (for people with multiple sclerosis),[1] but this property was untested in the other ordinal scales. When reported, the full range of scores was used in all the tools indicating a lack of floor or ceiling effects.[15–17,19,20,25–27] Criterion-related validity between measures of pain intensity was moderate to good, while concurrent validity between measures of pain intensity and pain interference, mood (or psychological functioning), and quality of life was weak to moderate.

Table 3. Summary of the validity of tools measuring pain in neurological conditions.

Scale	Construct measured	Validated for	Concurrent/criterion-related validity	Internal consistency and factor analysis	Construct validity including scaling properties (hierarchy/redundancy/ floor ceiling effects)
Faces Pain Scale-6 or 7[15–17]	Pain intensity	Adults with CP	Other measures of pain intensity +++; pain interference +; mood +	A valid measure of pain intensity	Not tested
	Shoulder pain intensity	Stroke	Other measures of pain intensity ++/+++; QoL +/++; mood +/++	Not tested	Poor construct validity. No RHBD and only 42% of LHBD could rank the faces for pain severity. Preferred measure for LHBD (over verbal/Likert scales or VAS). Full range of scores reported (no floor or ceiling effects).
NRS[15,16,19]	Pain intensity	Adults with CP	Other measures of pain intensity ++/+++; pain interference +/++; mood +	Not applicable	Not tested
Verbal Rating Scale/ Likert Scale[15–17]	Pain intensity	Adults with CP	Other measures of pain intensity ++/+++: pain interference +/++; mood +	Not applicable	Not tested
	Shoulder pain intensity	Stroke	Other measures of pain intensity ++/+++	Not applicable	Least preferred by RHBD or LHBD
VAS/NRS combined[15–17]	Shoulder pain intensity	Stroke	Other measures of pain intensity +++	Not applicable	Preferred method by RHBD. Full range of scores reported (no floor or ceiling effects)
VAS[15,17,19,20]	Shoulder pain intensity	Stroke	Other measures of pain intensity ++/+++	Not applicable	Full range of scores reported (no floor or ceiling effects
Pain assessment of cerebral palsy[21]	Intensity of hip pain	Adults with CP	Not tested	Internal consistency ++/+++	Pictures of painful activities scored higher than non-painful ones. Weak agreement between patients and caregivers scores
ShoulderQ[23,24]	Shoulder pain severity	Hemiplegic acquired brain injury	Not tested	Not tested	++ correlation between verbal and visual scoring methods (confirming different methods suited different impairments)

(Continued)

Table 3. (Continued)

Scale	Construct measured	Validated for	Concurrent/criterion-related validity	Internal consistency and factor analysis	Construct validity including scaling properties (hierarchy/redundancy/ floor ceiling effects)
Ritche Articular Index[20,22]	Shoulder pain intensity	Stroke	Other measures of pain intensity ++	Not tested	Not tested
Neuropathic Pain Scale[1]	Neuropathic pain severity	Multiple sclerosis	Other measures of pain intensity ++ mood +; QoL: body pain ++; MS severity (EDSS) +	3 factors accounted for 64% of variance. Internal consistency ++	Inter-item correlation +/++ (indicating no redundancy). Full range of scores used (no floor or ceiling effects). Most patients could define their pain using the descriptors
Pain Interference Scale of the Brief Pain Inventory[15,25,26]	How pain interferences with life	Spinal cord injuries; Adults with CP; multiple sclerosis	Spinal cord injuries: pain intensity ++; mental health ++. Multiple sclerosis: pain intensity +/++; mental health +/ ++. Adults with CP (10-item version): pain intensity ++; disability +	Multiple sclerosis: pain intensity and interference are separate constructs. Internal consistency Spinal Cord injuries +++. Multiple Sclerosis +++. Adults with CP (10 item version) +++	Spinal cord injuries: full range of scores used (no floor or ceiling effects).
Graded Chronic Pain Disability Score[18,25,26]	How pain interferences with life	Spinal cord injuries; Adults with CP; multiple sclerosis	Spinal cord injuries: pain intensity ++; mental health ++. Multiple sclerosis: pain intensity ++; mental health +. Adults with CP (10 items) pain intensity +; disability +	Multiple sclerosis: pain intensity and interference are a single factor. Internal consistency: spinal cord injuries +++; multiple sclerosis +++	Spinal cord injuries: full range of scores used (no floor or ceiling effects).

CP: cerebral palsy; EDSS: Expanded Disability Status Scale; LHBD: left hemisphere brain damage; MS: multiple sclerosis; NRS: Numerical Rating Scales; QoL: quality of life; RHBD: right hemisphere brain damage; VAS: visual analogue scale.

Considerable attention had been given to the way in which the items should be presented for best effect and differences between patient groups were noted. Using complex images, such as facial expressions, was unsuccessful as many patients, particularly those with right-hemisphere damage, struggled to recognize that the images represented pain (rather than sadness or anger) and few were able to rank them in the correct order of severity.[16,17] However the use of numeric rating or VASs was the preferred style for people with left hemisphere damage.[17] Both groups found rating scales using verbal or written descriptors least acceptable.[16,17]

Discussion

To the authors' knowledge this is the first attempt to systemically evaluate measures of pain in neurological conditions and so comparisons are limited. Despite the variety in the purpose and format of the selected tools, the review has revealed that none have been adequately developed and evaluated to recommend them for use in clinical practice or the research setting. All are straight forward and feasible to use but, as yet, it has not been demonstrated that they provide robust information.

All the measurement tools had undergone some examination of the concurrent and criterion-related validity and generally demonstrating a moderate correlation between measures of pain, and as would be expected, weak–moderate relationships with other constructs, such as mood and quality of life.

Content validity had received less consideration. Apart from the ShoulderQ,[23,24] the selected measurement tools were all originally developed for people with other painful conditions and then applied to neurological conditions. In doing so, it has been assumed that the content reflects neurological patients' experience of pain. This is logical when the construct being measured is pain intensity as, initiatively, this is a universal and paramount issue. However, it is notable that there are no reports of patients' involvement or consultation in the development or validation of these tools, and so it unknown to what extent they capture the issues that are important to the patients; other aspects of pain, such as the frequency or the nature of the pain, may be as high, or higher, priority. Patient consultations are needed to establish the content validity of the visual analogue and rating scales and the important aspects of pain that should be measured.

The extent to which the scales of pain interference are fit-for-purpose in neurological conditions is more uncertain. Although they have apparent face validity, the items involve evaluation of the pain's impact on everyday activities (such as mobility, sleep, and self-care) that are frequently affected by the neurological condition, independent of any pain. Consequently patients may struggle to identify whether activity limitations are due to pain or other impairments. Further work to explore neurological patients' experience of pain and its impact is a priority to ensure that measures reflect the problems that are the most relevant and important to them. Our findings of similar degrees of reliability and validity across groups gives some support to the assumption that the nature of pain and the associated difficulties generalize from one clinical group to another, and thus that psychometric properties for measurement tools could generalize between clinical conditions. Nevertheless, further work is needed to more comprehensively assess measurement tools across clinical groups and test this assumption.

The scaling properties of the ordinal scales were largely un-explored. There had been evaluation of floor and ceiling effects (which were found to be absent) in the Neuropathic Pain and pain interference scales,[15–17] but the Neuropathic Pain Scale[1] was the only tool in which redundancy of items had been assessed (and found to be absent). Other important aspects of scale construction, such as the hierarchy of items and fit to the Rasch (or similar) measurement model, have not been evaluated. This is an important omission as the scores for the multi-item scales are summated (for the Pain Interference Scale[18,25,26]) or an average taken (for the Graded Chronic Pain Disability Score[18,25,26]). This is an inappropriate use of the data from the measures as these calculations assume continuous (interval or ratio level) data, while the data produced are categorical (ordinal or nominal level). It is a relatively simple process to use Rasch analysis (or similar) to

assess and modify scale structure so that ordinal data can be converted to interval level and then summated (or composite/averaged) scores calculated and parametric statistics used. This analysis needs to be undertaken with some urgency, as the continued use of ordinal data as if it were continuous could lead to ineffective clinical decisions or inaccurate research conclusions being made. In addition, future publications should require that the full operating procedure or instruction manual is included or made easily available.

Reliability of the selected tools had received less attention than the validity, particularly inter-tester reliability, which was only assessed for two tools.[17,19] As pain is inherently subjective and its evaluation is dependent on self-report, one could argue that inter-tester reliability is not an issue. However, as many patients with neurological conditions have difficulty completing standardized measurement tools, it is common practice for others (staff, family, or carers) to assist them, and the way in which they do this is likely to impact on the scores obtained. This hypothesis is supported by Pomeroy et al.[19] who reported poor inter-tester reliability (once the scores for people with no pain were removed) when using VASs in a clinical setting, which they attributed to lack of standardization and assessor training. If these methods were used in clinical practice, they could cause ineffective, or possibly harmful, clinical decisions to be made. More thorough assessment of test–retest and inter-tester reliability and ability to detect change is needed to address this issue. More thorough operating (and scoring) instructions are also needed to establish how patients should complete the scales and how others should/could help them.

Measuring pain in people with neurological conditions is a challenge as its inherently subjective nature means evaluation tools rely on self-report, but confounding impairments make this difficult for many patients. Cognitive, communication, and/or visual impairments can limit patients' ability to understand the questions, transfer their experience to a score, or express their answer. Alternative ways of obtaining scores, such as reports from caregivers or scores based on patients' behavioural responses, show poor agreement with self-report and open to misinterpretation.[19,21] Thus, self-report of pain would appear the most effective design to use and the imperative is to establish the most effective way to elicit an accurate report. However, no single format suits all patients. People with left hemisphere damage prefer and report most reliably when using visual images or numbers (but not words), while those with right hemisphere damage do better when reporting their pain using simple images or numbers (but not written or verbal responses).[17] Both stroke survivors and adults with cerebral palsy have difficulty recognizing that complex images, such as facial expression, represent pain and could not rank them in order of severity.[16,17] Thus, the Facial Pain Scale should not be used. Overall, numerical rating or VASs were the most effective.[16,17] Although there is no evidence to demonstrate a difference, it has become accepted that a vertical rating scale is preferable to a horizontal one as it reduces the challenge for people with visual, scanning, or attention deficits; a colour grading may also be helpful.[15–17,19,23,24] However, further work is needed to identify the optimal presentation(s) and method(s) of data collection across a range of neurological conditions and impairments.

Study limitations

The quality of this review is dependent on the articles identified. Although we had thorough search strategies, we only included publications in English. There may have been relevant publications in other languages that we missed. We also did not attempt to identify unpublished data or the grey literature, so there may have been a publication bias in the data that we identified. To our knowledge, this is the first systematic review to specifically assess the clinical utility of measures of pain in neurological conditions. The system we developed to assess the utility was based on our clinical experience, with consensus from neurological physiotherapists and the judgements of quality were arbitrary. Although they have strong face validity, such judgements cannot be assumed to be appropriate for other healthcare systems or other areas of clinical practice.

Clinical messages

- The selected pain measurement tools were feasible to use, but none demonstrated sufficient psychometric properties to recommend.
- Numeric rating scales or VASs appear to suit most patients, while written rating scales seem least effective. People with right hemisphere damage struggle with images and written layouts, while people with left hemisphere damage struggle with written and verbal layouts.
- Measures involving complex images, such as facial expression, should not be used as patients find them difficult to interpret and rank.

Conflict of interest

The author declares that there is no conflict of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Rog DJ, Nurmikko TJ, Friede T and Young C. Validation and reliability of the Neuropathic Pain Scale in multiple sclerosis'. *Clin J Pain* 2007; 23(6): 473–481.
2. Archibald CJ, McGrath PJ, Ritvo PG, et al. Pain prevalence, severity and impact in a clinic sample of multiple sclerosis patients. *Pain* 1994; 58(1): 89e93.
3. Solaro C, Brichetto G, Amato MP, et al. The prevalence of pain in multiple sclerosis: a multicenter cross-sectional study. *Neurology* 2004; 63(5): 919e921.
4. Ehde DM, Gibbons LE, Chwastiak L, et al. Chronic pain in a large community sample of persons with multiple sclerosis. *Mult Scler* 2003; 9(6): 605e611.
5. Ehde DW, Jensen MP, Engel JM, et al. Chronic pain secondary to disability: A review. *Clin J Pain* 2003; 19: 3–17.
6. Schwartz L, Engel JM and Jensen MP. Pain in persons with cerebral palsy. *Arch Phys Med Rehabil* 1999; 80: 1243–1246.
7. Jönsson A-C, Lindgren I, Hallström B, et al. Prevalence and intensity of pain after stroke: a population based study focusing on patients' perspectives. *J Neurol Neurosurg Psych* 2006; 77: 590–595.
8. Broen MP, Braaksma MM, Patijn J and Weber WE. Prevalence of pain in Parkinson's disease: a systematic review using the modified QUADAS tool. *Mov Disord* 2012; 27(4): 480–484.
9. Oliver D. The quality of care and symptom control—the effects on the terminal phase of ALS/MND. *J Neurol Sci* 1996; 139: 134–136.
10. Ganzini L, Johnston WS and Hoffman WF. Correlates of suffering in amyotrophic lateral sclerosis. *Neurology* 1999; 52(7): 1434–1440.
11. Newrick PG and Langton-Hewer R. Pain in motor neuron disease. *J Neurol Neurosurg Psych* 1985; 48(8): 838–840.
12. Tyson S, Watson A, Moss S, et al. Development of an evidence based framework for the physiotherapy assessment of neurological conditions? *Disab Rehabil* 2007; 30(2): 142–144.
13. Connell L and Tyson SF. Measures of sensation in neurological conditions: a systematic review. *Clin Rehab* 2012; 26(1): 68–80.
14. Tyson S and Connell LA. How to measure balance activity in clinical practice? A systematic review of the psychometric properties and clinical utility of measurement tools in neurological conditions. *Clin Rehab* 2009; 23: 824–840.
15. Dogan SK, Ay S, Oztunda D, et al. The utility of the Faces Pain Scale in the assessment of shoulder pain in Turkish stroke patients: its relation with quality of life and psychologic status. *Int J Rehabil Res* 2010; 33(4): 363–367.
16. Jensen MP, Engel JM, McKearnan KA and Hoffman AJ. Validity of pain assessment in persons with cerebral palsy: a comparison of six scales. *J Pain* 2003; 4(2): 56–63.
17. Benaim C, Froger J, Cazottes C, et al. Use of the Faces Pain Scale by left and right hemispheric stroke patients. *Pain* 2007; 128: 52–58.
18. Tyler EJ, Jensen MP, Engel JM, et al. The reliability and validity of pain interference measures in persons with cerebral palsy. *Arch Phys Med Rehabil* 2002; 83: 236–239.
19. Pomeroy VM, Frames C, Fragher EB, et al. Reliability of a measure of post-stroke shoulder pain in patients with and without aphasia and/or unilateral spatial neglect'. *Clin Rehab* 2000; 14: 584–591.
20. Gustaffson L and McKenna K. Correlation between two measurement scales of hemiplegic shoulder pain. *Int J Ther Rehabil* 2003; 10: 247–252.
21. Bouldingh EJ, Jacobs-van der Bruggen MA, Lankhorst GJ et al. Assessing pain in patients with severe cerebral palsy: development, reliability and validity of a pain assessment instrument for cerebral palsy. *Arch Phys Med Rehabil* 2004; 85: 758–766.
22. Bohannon R and LeFort A. Hemiplegic shoulder pain measured with the Ritchie Articular Index. *Int J Rehabil Res* 1986; 9(4): 279–281.

23. Turner-Stokes L and Jackson D. Assessment of shoulder pain in hemiplegia: sensitivity of the ShoulderQ. *Disab Rehab* 2006; 28(6): 389–395.
24. Turner-Stokes L and Rusconi S. Screening for ability to complete a questionnaire: a preliminary evaluation of the AbilityQ and ShoulderQ for assessing shoulder pain in stroke. *Clin Rehabil* 2003; 17: 150–157.
25. Osborne TL, Raichle KA, Jensen MP, et al. The reliability and validity of pain interference measures in persons with multiple sclerosis. *J Pain Symp Manag* 2006; 32: 217e229.
26. Raichle KA, Osborne TL, Jensen MP, et al. The reliability and validity of pain interference measures in persons with spinal cord injury. *J Pain* 2006; 7(3): 179–186.

Appendix 1. Details of the data extracted from the selected articles reporting psychometric properties of measures of pain in neurological conditions.

Authors and date	Psychometric property tested	Subjects	Procedure (what was done)	Analysis	Results
Bouldingh et al.[21]	Test–retest reliability; internal consistency; validity of patients' scores against caregivers scores of the PAICP	Children and adults with CP and hip problems. Reliability study 4 severely disabled children with CP. Validity study = 160 adults with severe CP. 87 men; mean age = 36 years (16–84). 19 were non-verbal and 11 used communication aids	All participants completed the PAICP with a researcher. Participants in the reliability study completed the assessment again 2 weeks later.	Test–retest reliability: Kappa and percentage agreement. Validity: between patients and caregivers; Spearman correlation Internal consistency: Cronbach's alpha	Test–retest reliability (moderate–excellent): k = 0.48–1.0. Percentage agreement = 70–100% Internal consistency (good-excellent): α = 0.65–0.83. Construct validity (mixed): Caregivers' scores were the same as patients in 88% of observations. Correlations between staff and patients' scores were poor. Physio vs. patient; *r* = 0.02–0.52. Caregivers vs. patient; *r* = 0.01–0.48
Dogan et al.[15]	Validity of vertical Faces Pain Scale-7 (FPS-7-v) vs. horizontal 10 cm VAS; LPS; 11-point NPRS-11 vs. SF-36 bodily pain sub-scale; QoL (full SF-36); mood (BDI)	30 strokes with shoulder pain. Mean age: 64 ±9 years; duration of stroke-related pain = 6 (±3) months; 16 males; 21 left hemiplegia. Excluded people with cognitive defects	Participants completed all the measures to rate their shoulder pain. No details about how delivered	Validity: Spearman's correlation	Criterion- related validity (excellent): FPS-7 vs. VAS *r* = 0.950; LPS *r* = 0.972; NPRS-11 *r* = 0.957; SF-36 (pain) *r* = 0.707. Concurrent validity (weak– moderate): FPS-7 vs. SF-36 *r* = 0.128–0.461; BDI = 0.29
Jensen et al.[16]	Criterion-related validity of different rating scales: 11- and 21-point NRS; 5- and 16-point VRS; and 6- and 7-point FPS. Concurrent validity with mood. Pain interference in daily life (PIS-BPI)	69 adults with CP and chronic pain. Mean age: 40 years (±13); 54% male. 58% spastic, 13% athetoid, 1% ataxic, and 25% mixed. All had chronic pain. Excluded if IQ < 70.	All measures completed with an interviewer one after other. 24 participants did not complete the NPRS-11. Only 45 completed the assessment of mood and pain interference	Concurrent and criterion validity: correlation co-efficients. Construct validity: principal components analysis of the ratings scales	Criterion-related validity (good–excellent): FPS-7 vs. NRS-11 *r* = 0.59; NRS-11 vs. NRS-21 *r* = 0.87. NRS-21 had strongest correlation with the other measures (*r* = 081–0.84). Concurrent validity (weak–good): with measures of PIS-BPI vs. NRS-11 *r* = 0.25; FPS-7 *r* = 0.7. Weak with mood *r* = 0.23 (VRS-5, worst) to *r* = 0.38 (FPS-7, best). Construct Validity: single factor accounted for 82–86% of variance.

(Continued)

Appendix 1. (Continued)

Authors and date	Psychometric property tested	Subjects	Procedure (what was done)	Analysis	Results
Turner-Stokes and Jackson[23]	Responsiveness of the ShoulderQ by assessing whether changes to ShoulderQ is in keeping with clinical response. Sensitivity and specificity to detect likely responders to pain management pathway	30 adults with hemiplegic shoulder pain and severe complex acquired brain injury. Mean age: 47 years (±2); 50% left hemplegia and 10% bilateral.	ShoulderQ completed fortnightly. On discharge clinical team rated whether the patient had responded (n = 18) or not (n = 12)	Paired t-tests assessed change in scores with time. Sensitivity and specificity (+ve and –ve predictive value)	Improvements in pain at night, on movement, total, and verbal rating score in responders but not non-responders. Change in score of 3+ has 77% sensitivity and 91% specificity to identify responders; +ve predictive value of 93%. A score of 2+ had -ve predictive value of 73%. Zero or minus changes have -ve predictive value of 60%.
Turner-Stokes and Rusconi[24]	Test–retest reliability of the ShoulderQ.	49 severe strokes (mean age 53 years (±3), female/male ratio 2:1. Mean time since stroke 4 months (±3). 57% had communication deficits; 63% had cognitive deficits	22/49 participants with shoulder pain completed the ShoulderQ on consecutive days with assistance as necessary.	Test–retest reliability: percentage agreement for 'yes/no' questions. Kappa statistics for Likert scales VAS scores = ICCs	Test–retest reliability (moderate): percentage agreement = 36–100%; k = 0.16–0.60; ICCs = 0.50–0.60. Poorer validity scores for Qs regarding pain frequency and severity during physio.
Pomeroy et al.[19]	Test–retest reliability and inter-rater reliability of three 10 cm vertical-VAS to assess pain severity, frequency, and bothersomeness. Also question about the nature of the pain	33 strokes with shoulder pain. Mean age: 74 years (57–89); 14 men; mean time since stroke 42 months (7–360). 8 were aphasic; 8 had neglect; 5 had both.	Tests administered by physiotherapists on 3 consecutive days (within 2 hours on each day). NB. Physiotherapists assessed by observation. Tests were performed without standardized operating procedures or training before administration	For the VASs: test–retest and inter-rater reliability = ICC. For the yes/no nature of pain Q; Kappa and percentage agreement. 95% LoA (mean difference ±2)	Test–retest reliability (moderate–good): severity of pain ICCs = 0.75–0.81. Frequency of pain: ICCs = 0.74–0.80. Bothersomeness. ICCs = 0.55–0.72. Nature of pain Q = 35–66% agreement. Inter-rated reliability (poor-good): severity of pain ICCs = 0.64–0.74. Frequency of pain ICCs = 0.71–0.88. Bothersomeness. ICCs = 0.64–0.77. Nature of pain Q = 48–76% agreement. Most agreement was from zero scores. When pain was present, wide variability between raters seen. Assessment of pain severity and frequency was more reliable than bothersomeness.

Appendix 1. (Continued)

Authors and date	Psychometric property tested	Subjects	Procedure (what was done)	Analysis	Results
Gustaffson and McKenna[20]	Criterion-related validity of 10 cm vertical-VAS (severity of shoulder pain) vs. RAI.	34 strokes with shoulder weakness. Mean age: 66 years (±14). 20 women. Mean 24 days post stroke (±25). 14 had right hemiplegia	Participants rated severity of their pain using the VAS and assessor performed the RAI on their affected shoulder.	Spearman's correlation coefficient.	Concurrent validity (moderate): r = 0.57.
Rog et al.[1]	Validity of NPS vs. Short Form McGill Pain Questionnaire (SF-McGill), SF-36 (bodily pain item); full SF-36; HADS. Construct validity, internal consistency, factor analysis, validity of postal vs. hospital completion, test–retest reliability of the NPS, measurement error (95% LoA)	129 patients with MS. Median age: 42 years; 84% female. 52% relapse remitting; median duration = 6.5 years	All participants completed the NPS, SF-McGill, and HADs, SF-36 in clinic, day ward, or by post. They were unsupervised during completion. A scribe assisted as needed. 116 repeated the NPS 4–5 weeks later (analysis only on patients who felt their pain had not changed). 71 participants completed the NPS at home and then in clinic or the day- unit within a week (to assess whether NPS could be used by post)	Internal consistency: Cronbach's alpha. Factor analysis: principal components with varimax factor rotation. Validity: Spearman correlation co-efficient. Test–retest reliability: ICCs. Agreement between postal and clinic completion 95% LoA (Bland Altman plots)	Internal consistency (good): α = 0.78. Factor analysis: 3 factors explained 64% of variance; Factor 1 = intense, sharp, unpleasant, and deep pain. Factor 2 = sensitivity, itchy, and surface pain. Factor 3 = dull and cold pain. Criterion related validity (weak–moderate): SF-McGill r = 0.49–0.63; SF-36 (pain) r = –0.49. Concurrent validity (weak) vs. QoL (SF-36) r = –0.08 to –0.49 vs. mood (HADS) r = 0.22–0.27. Test–retest reliability (moderate–good): ICC (total NPS) = 0.71; individual items ICC r = 0.32 (surface), r = 0.84 (sensitivity). Agreement between postal and clinic completion (n = 71) 95% LoA = ±14 points

(Continued)

Appendix 1. (Continued)

Authors and date	Psychometric property tested	Subjects	Procedure (what was done)	Analysis	Results
Benaim et al.[17]	Validity of the FPS-7 to assess shoulder pain. Criterion-related validity vs. the VAS-vertical (10 cm) and VRS (5-point likert scale). Test–retest and inter-rater reliability of the FPS-7. Rank order analysis of the FPS-7.	127 strokes. Mean age = 63 years (±8), 35% female. Mean time since stroke = 69 days. 50% = left hemiplegia. In the right hemiplegics; 60% had aphasia and 59% had unilateral spatial neglect. 51 took part in the validity study. 33 took part in the test–retest reliability study 43 took part in the inter-rater reliability study	Participants completed each assessment tool in a random order. Those in the test–retest reliability study repeated the assessment 1–2 days later. The assessment was also completed by another assessor 1–2 days later to asses inter-rater reliability. Content validity: patients ranked the faces on the FPS-7 to assess whether they recognized the faces represented pain severity.	Content validity- Rank order of the pictures. (% of participants that correctly ordered all the faces). Concurrent validity Spearman correlation between v-FPS-7 and v-VAS and VRS. Reliability Kappa and ICC's for VAS. All data analysed separately for KHBD and LHBD	Content validity: 48% of RHBD and 84% of LHBD recognized the faces represented pain. RHBD tended to think the face represented sadness > pain. Ability to rank the faces was poor. None of the RHBD could do so correctly only 42% of LHBD. Of the 71 who could express a preference, most LHBD preferred the FPS-7 (16/32) and most RHBD preferred the VAS (18/39). Criterion-related validity (moderate-excellent): FPS-7 vs. VRS r = 0.65–072 (RHBD > LHBD).VAS r = 0.72–0.82 (LHBD > RHBD). Test–retest reliability (weak–excellent): FPS k = 0.53–0.74 LHBD > RHBD). VAS ICC = 0.72–0.86 (RHBD > LHBD). VRS k = 0.39–0.57 (RHBD > LHBD). Inter-rater reliability (moderate–excellent): FPS k = 0.44–0.64 (LHBD > RHBD). VAS ICC = 0.72–0.86. (RHBD > LHBD). VRS k = 0.46, k = 0.52. (RHBD > LHBD).
Bohannon and LeFort[22]	Inter-rater reliability of the RAI to assess hemiplegic shoulder pain.	34 strokes. Mean age: 65 years (±10); 8 women. Mean time since stroke 98 days (±217).	2 raters observed patient's behaviour while their shoulder was externally rotated passively using a standardized procedure. The 2nd rater tested each patient a few minutes later	Inter-rater reliability: Kappa and percentage agreement	The raters' scores agreed on 85% of assessments and never exceeded 1 point difference. 1 rater diagnosed that 27/34 patients had shoulder pain and the other diagnosed 26/34 cases. Inter-rater reliability: k = 0.759

Appendix 1. (Continued)

Authors and date	Psychometric property tested	Subjects	Procedure (what was done)	Analysis	Results
Tyler et al.[18]	Validity of pain interference on everyday life: the CPG and PIS-BPI vs. pain intensity; global disability (CHART and mood CES-D Scale)	50 adults with CP, pain, and an IQ >70; 25 women. Mean age = 40 years (sd 13). 84% had a quadriplegia; 58% had spastic CP, 26% had mixed CP, 14% CP, and 2% hypotonic-type CP. Most were non-ambulatory (80%)	All measures rated in a patient interview. NRS pain intensity = average pain over the last 24 hours (0–10); CPG = how their pain's interfered with ADL; recreational, social, and family activities; work/ housework (0–10 and the scores averaged)	Validity: correlation co-efficients. Internal consistency: Cronbach's Alpha.	CPG Concurrent validity (weak): Pain intensity. r = 0.35; CES-D score r = 0.45; CHART r = 0.16. Internal consistency (moderate): α = 0.59. PIS-BPI Validity (moderate–good): pain intensity (r = 0.66); CES-D score (r = 0.45); CHART (not reported). Internal consistency (excellent): α = 0.89
Osborne et al.[25]	Construct, content and criterion-related validity of two measures of pain interference on everyday life: the CPG and PIS-BPI	125 with MS-related pain. 75% women; mean age = 51 years (±11). 53% had relapsing/remitting MS. Mean duration = 13 years (±10). 46% reported mild pain in the last week; 27% moderate pain, 25% had severe pain	Average and worst pain in the past week rated with a (0–10) NRS and compared with the CPG and 3 versions of the PIS-BPI (7-, 10- and 12-item versions). 10-item version has additional questions about the impact on daily life, activity, and participation, and the 12-item version included impact on communication and cognition	Internal consistency: Cronbach's Alpha. Validity: Pearson correlation. Construct validity: factor analysis.	Internal consistency (excellent): for PIS-BPI; 7-item = 0.93; 10-item = 0.95; 12-item = 0.96 GCP = 0.94. Factor analyses: pain interference and intensity items were separate factors accounting for ~60% and ~10%. of variance respectively. Validity (moderate) vs. pain intensity NRS vs. PIS-BPI-7 r = 0.63; PIS-BPI-10 r = 0.63; PIS-BPI-12 r = 0.61; GCP r = 0.62–0.67. Weak vs. SF-Mental Health; PIS-BPI-7 r = 0.47; PIS-BPI-10 r = 0.48; PIS-BPI-12 r = 0.51; GCP r = 0.25–0.34

(Continued)

Appendix 1. (Continued)

Authors and date	Psychometric property tested	Subjects	Procedure (what was done)	Analysis	Results
Raichle et al.[26]	Construct, content, and criterion-related validity of two measures of pain interference on everyday life: the CPG and 3 versions of the PIS-BPI	127 with spinal cord injury and pain. 72% men; mean age = 49 years (±13); mean time since SCI = 17 years (±11). 58% quadriplegic.	Participants completed the CPG and 3 versions of the PIS-BPI (7-, 10- and 12-item versions), plus a NRS (0–10) of the average pain intensity over the past week. And psychological function using the MH-SF36	Internal consistency: Cronbach's Alpha. Validity: Pearson correlation with MH-SF36). Construct validity: factor analysis of the PIS-BPI-7 and the Pain Intensity NRS	Internal consistency (excellent): α = PIS-BPI = 0.95–0.96 (all versions) CPG = 0.86–0.95. Validity: PIS-BPI vs. Pain Intensity NRS = good PIS-BPI-7 r = 0.62; PIS-BPI-10 r = 0.63; PIS-BPI-12 r = 0.61; vs. MH-SF36; PIS-BPI-7 r = 0.62; PIS-BPI-10 r = 0.60; PIS-BPI-12 r = 0.61; moderate for the CPG = NRS r = 0.5; MH-SF36 r = 0.55

ADL: activities of daily living; BDI: Becks Depression Inventory; CES-D: Centre for Epidemiological Studies Depression Scale; CHART: Craig Handicap Assessment and Reporting Technique; CP: cerebral palsy; CPG: Chronic Pain Grade; FPS: Faces Pain Scale; HADS: Hospital Anxiety and Depression Scale; ICC: intraclass correlation coefficient; IQ: Intelligence Quotient; LHBD: left hemisphere brain damage; LoA: Limits of Agreement; LPS: Likert Pain Scale; MH-sF: Mental Health scale from the SF; MS: multiple sclerosis; NPRS: Numeric Pain Rating Scale; NPS: Neuropathic Pain Scale; NRS: Numerical Rating Scale; PAICP: Pain Assessment Instrument of Cerebral Palsy; PIS-BPI: Pain Interference Scale of the Brief Pain Inventory; QoL: quality of life; RAI: Ritchie Articular Index; RHBD: right hemisphere brain damage; SF: Short Form; SF-McGill: Short Form McGill Pain Questionnaire; VAS: visual analogue scale; VRS: Verbal Rating Scale.