Pain and Physical Functioning in Neuropathic Pain: A Systematic Review of Psychometric Properties of Various Outcome Measures

Poonam Mehta, MPT*; Leica S. Claydon, PhD†; Paul Hendrick, PhD‡; Chad Cook, PhD§; David G. Baxter, Dphil*

*Centre for Physiotherapy Research, University of Otago, Dunedin, New Zealand; †Department of Allied Health and Medicine, Anglia Ruskin University, Chelmsford; ‡Division of Physiotherapy Education, University of Nottingham, Nottingham, U.K.; §Division of Physical Therapy, Walsh University, North Canton, Ohio, U.S.A.

Abstract

Introduction: A range of outcome measures across various domains are used to evaluate change following an intervention in clinical trials on chronic neuropathic pain (NeP). However, to capture a real change in the variable of interest, the psychometric properties of a particular measure should demonstrate appropriate methodological quality. Various outcome measures in the domains of pain and physical functioning have been used in the literature for NeP, for which individual properties (eg, reliability/validity) have been reported. To date, there is no definitive synthesis of evidence on the psychometric properties of those outcome measures; thus, the aim of this systematic review was to evaluate the methodological quality [COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) guidelines] of studies that evaluated psychometric properties of pain and physical functioning outcome measures used for NeP.

Methods: Specific MeSH/keywords related to 3 areas (pain and/or physical functioning, psychometric properties, and NeP) were used to retrieve relevant studies (English language) in key electronic databases (MEDLINE (Ovid), CINAHL (EBSCO), Scopus, AMED, and Web of Science) from database inception—July 2012. Articles retrieval/screening and quality analysis (COSMIN) were carried out by 2 independent reviewers.

Results: Twenty-four pain and thirty-seven physical functioning outcome measures were identified, varying in methodological quality from poor–excellent.

Conclusion: Although a variety of pain and physical functioning outcome measures have been reported in the literature, few have demonstrate methodologically strong psychometric properties. Thus, future research is required to further investigate the psychometric properties of existing pain and physical functioning outcome measures used for clinical and research purposes.

Key Words: neuropathic pain, systematic review, pain, physical function, outcome measures, psychometric properties, reliability, validity, responsiveness

INTRODUCTION

Neuropathic pain (NeP) is defined by the International Association for the Study of Pain’s Neuropathic Pain Special Interest Group (NeuPSIG) as “pain arising as a direct consequence of a lesion or disease affecting the...
A range of assessment guidelines have been developed from the Initiative on Methods, Measurement and Pain Assessment in Clinical Trials (IMMPACT), the European Federation of Neurological Societies (EFNS), and the NeuPSIG for NeP clinical trials and for clinical practice. These guidelines advocate a range of measures for assessing the core domains of pain, quality of life, mood, sleep, and functional capacity (physical, cognitive, emotional, and social). This notwithstanding, a variety of outcome measures are available for the above-stated domains. To evaluate the applicability of these measures, a systematic review of psychometric properties of available outcome measures used in published trials may provide a useful basis for selecting the best measurement instrument for a specific purpose.5,6

Individual assessment of psychometric properties of available outcome measures is important.7 As part of this, in reviewing the evidence on available outcome measures, it is important to assess the methodological quality of those studies that investigated psychometric properties.9 While in clinical practice, adoption of outcome measures will depend on feasibility of use (speed, ease of use, and limited need for an overly sophisticated instrument),10 emphases should also be given to measures which are proven to be reliable, valid, and responsive/interpretable for a given population.

Pain remains a leading cause of disability at the individual level, associated with functional losses as well as mood disturbances.11 Thus, the focus of this systematic review will be in evaluating the psychometric properties of various outcome measures used in the domains of pain and physical functioning in NeP. On examination of the literature, a number of outcome measures have been identified in which have been used to measure pain intensity and physical function in NeP trials;5,7,8,12 however, there is limited conclusive evidence on their psychometric properties. Use of reliable and valid outcome measures can help to better evaluate the patient’s outcomes in terms of pain and physical functioning, enabling better management, including the earliest appropriate management to minimize risks of comorbidities and disabilities.

Existing evidence on the psychometric properties of pain and physical functioning outcome measures used in NeP trials have not previously been systematically reviewed. The aim of this systematic review was to systematically review and identify the gaps in literature for the evaluated psychometric properties (reliability, validity, responsiveness, and interpretability) of identified outcome measures for “pain and physical functioning” as recommended by the IMMPACT guidelines in NeP population. This review involved a systematic search of the literature. The findings of the current study may assist in outlining the effective intervention strategies for patients with NeP. The objectives of this systematic review were as follows: (1) Systematically review and identify the type of established psychometric properties for the identified outcome measures quantifying pain and physical functioning in neuropathic pain populations; (2) Evaluate the methodological quality of the included studies investigating the psychometric properties of the identified outcome measures in the domain of pain and physical functioning in neuropathic pain populations in accordance with the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) checklist with 4-point scale.

**METHODS**

**Information Sources**

A systematic search was conducted following the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines. The following electronic databases were searched: Ovid MEDLINE, CINAHL, Scopus, AMED, and Web of Science (WOS) (from database inception to July 31, 2012). The search update engine from the available databases was activated to be familiar with the new searches in the current field, since the original search.

**Search Strategy**

The keywords and MESH headings in 3 broader areas (pain and/or physical functioning outcome measures, psychometric properties, and NeP) were used in the development of a search strategy (Table 1). Several strategies were used to develop a comprehensive list of keywords/Mesh terms/subject headings representing each area. For outcome measures, all pain and physical functioning outcome measures that were used in clinical trials of NeP were chosen. For psychometric properties, we chose the standardized terminologies used by the COSMIN framework.6 For the terms relating to NeP, MESH terms/keywords indexed for neuropathy, neuralgia, and neurodynia were used. Words within each
theme were combined with OR and across themes with AND. This search strategy was amended for different databases as necessary.

\[\text{Study Selection}\]

Articles identified in the search underwent a series of screening processes. Firstly, duplicate articles were removed. Two reviewers (PM and LC) independently selected and screened articles for potential eligibility at the title and abstract stages. Full-text articles of all potentially eligible abstracts were retrieved for application of the eligibility criteria. Disagreements between the reviewers regarding inclusion of individual studies were discussed during a consensus meeting and, when unresolved, were resolved by discussion with other reviewers (PH, CC, and GDB). References of the selected papers were further explored for relevant articles.

\[\text{Eligibility Criteria}\]

Cross-sectional studies and longitudinal cohort studies, which included at least 1 assessment of a psychometric property of a pain or functional outcome measure in a NeP population (NeP as defined by the Clinical Resource Efficiency Support Team—CREST)\(^\text{13}\), were included. The adopted search strategy revealed 2 distinct categories of evaluations: one intended for screening or diagnosis; the other developed to measure outcomes. As the focus of this review was to investigate the psychometric properties of tools used to measure changes in the status of either pain or functional outcomes over time, screening or diagnostic tools were excluded. Studies published as case report, editorial, or reviews were also excluded. Only articles published in the English language and on humans were selected.

\[\text{Data extraction and Synthesis}\]

A systematic approach to data extraction was carried out by independent reviewers (PM and LC/PH/CC/GDB), with equal number of articles randomly distributed among the team members. Each member extracted the data from the allotted articles, which were then checked for accuracy, with consensus meetings and opinions from other reviewers to resolve any disagreements. The following data were collected and tabulated from each of the included articles: study reference, participant characteristics, outcome measures studied, and type of psychometric properties tested (reliability and/or validity) (Table 2). Further summary of identified outcome measures with their published psychometric properties and COSMIN grading were synthesized (Tables S1 & S2). Results from excellent and good methodological quality studies based on COSMIN criteria (as stated in Table S3) were used to formulate recommendations for acceptable psychometric properties scores (for definitions of acceptable, good and excellent scores see Table S3).

\[\text{Methodological Quality of Individual Studies Reporting on Psychometric Properties}\]

Whereas a variety of tools are available to measure the methodological quality of studies that report on scale development and assessed psychometric properties, the CO\(\text{N}\)sensus-based Standards for the selection

| Table 1. Search Strategy |
|--------------------------|
| **Theme 1 Psychometric Properties** | **Theme 2 Pain and/or Physical Functional Outcome Measures** | **Theme 3 Neuropathic Pain** |
| Clinometric properties OR Visual analog scale OR | Pain OR |
| Validity OR Numerical Pain Rating scale OR Nerve pain OR |
| Reliability OR McGill pain rating scale OR Neuralgia OR |
| Sensitivity OR Pain disability index OR Neurodynia OR |
| Responsiveness OR Functional component of The Western Neuropathy |
| Minimal(ly) clinically important difference OR Ontario and McMaster Universities Arthritis Index OR |
| Minimal(ly) clinically important change OR Timed scored functional activity OR |
| Minimum detectable change OR Functional reach test OR |
| Smallest detectable change Disability of the arm shoulder and hand questionnaire OR |
| | Ulnar Neuropathy at Elbow questionnaire OR |
| | Daily activities by Verbal Rating Scale OR |
| | Function interference by Numerical Rating Scale |
Table 2. Summary of Included Studies

| Reference | Participant's Characteristics | Outcome Measures Studied | Psychometric Properties Tested |
|-----------|------------------------------|--------------------------|-------------------------------|
| Alderson & McGall<sup>46</sup> | Carpal Tunnel Syndrome | n = 17 | Alderson-McGall hand function questionnaire | Reliability—Internal consistency, test-retest reliability; Validity—Convergent validity; |
| | | | | Validity—Discriminative validity |
| Amirjani et al.<sup>57</sup> | Carpal Tunnel Syndrome | n = 162 | Dellon-modified Moberg pick-up test | Reliability—test-retest reliability; Validity—Convergent validity; |
| | | | Modified Neuropathy Disability Score | Validity—Criterion validity; |
| Asad et al.<sup>63</sup> | Type 2 diabetes sensorimotor NeP | n = 60 | Gender | Reliability—Internal consistency, test-retest reliability; |
| Bastyr et al.<sup>28</sup> | Diabetic peripheral NeP | n = 205 | Gender | Validity—Construct and Convergent validity, Responsiveness |
| Bouhassira et al.<sup>25</sup> | Peripheral and Central NeP | n = 176 | Gender | Validity—Construct validity; |
| | | | | Validity—Criterion validity; |
| | | | | Validity—Hypothesis testing; |
| | | | | Responsiveness |
| Bril & Perkins<sup>39</sup> | Type 1 and 2 diabetes NeP | n = 89 | Gender | Validity—Construct validity; |
| Bril et al.<sup>21</sup> | Diabetic sensorimotor poly NeP | n = 65 | Gender | Reliability—Internal consistency, inter-rater, intrarater reliability; |
| Collins et al.<sup>41</sup> | Complex regional pain syndrome-1 | n = 27 | Gender | Reliability—test-retest reliability and measurement error |
| Coplan et al.<sup>44</sup> | Herpes Zoster | n = 121 | Gender | Validity—test-retest reliability; |
| Cornblath et al.<sup>40</sup> | Diabetic poly NeP | n = 30 | Gender | Validity—Hypothesis testing; |
| Crawford et al.<sup>26</sup> | Neuropathic Pain | n = 130 | Gender | Validity—Construct validity; |
| Davidoff et al.<sup>43</sup> | Reflex Sympathetic Dystrophy Syndrome | n = 17 | Gender | Validity—Construct validity; |
| de Andrade et al.<sup>32</sup> | Neuropathic Pain | n = 94 | Gender | Reliability—test-retest reliability; |
| Dias et al.<sup>53</sup> | Wrist and hand disorders due to nerve involvement | n = 26 | Gender | Validity—Face validity and Construct validity, Responsiveness |
| Dworkin et al.<sup>37</sup> | Diverse chronic pain syndrome; Diabetic NeP | n = 1108 | Gender | Validity—Construct validity, Responsiveness |
| Eklund et al.<sup>54</sup> | Charcot-Marie-Tooth disease | n = 20 | Gender | Validity—Hypothesis testing |
| Erdmann et al.<sup>48</sup> | Chronic idiopathic demyelinating polyneuropathy; Multifocal Mono neuropathy | n = 30 | Gender | Validity—Hypothesis testing |
| Farrar et al.<sup>27</sup> | Diabetic peripheral NeP; Fibromyalgia syndrome | n = 1700 | Gender | Validity—Hypothesis testing |
| Farrar et al.<sup>29</sup> | Diabetic peripheral NeP; Patherpetic Neuralgia | n = 984 | Gender | Validity—Construct validity; |
| Farrell et al.<sup>60</sup> | Patherpetic Neuralgia | n = 31 | Gender | Validity—Construct validity; |
| | | | | Responsiveness |
| | | | | Reliability—Internal consistency; |
| | | | | Validity—Criteriar validity and Hypothesis testing, Responsiveness |
| Reference | Participant’s Characteristics | n | Outcome Measures Studied | Psychometric Properties Tested |
|-----------|------------------------------|---|--------------------------|------------------------------|
| Felix & Widerström-Noga³³ | NeP related to Spinal Cord Injury | n = 22 | Quantitative Sensory Testing (cold and heat pain thresholds) | Reliability—inter-rater and test-retest reliability; Validity—Construct validity |
| Galer & Jensen²⁵ | Postherpetic Neuralgia; Diabetic NeP; Peripheral Nerve Injury | n = 160 (69; 24; 67) | The Neuropathic Pain Scale | Validity—Hypothesis testing—Discriminative validity and Predictive validity |
| Geber et al.³³ | Peripheral Nerve lesion; Other neuropathies | n = 60 | Quantitative Sensory Testing (heat, cold, mechanical, and pressure pain threshold) | Reliability—inter-rater and test-retest reliability |
| Graham & Hughes⁷⁷ | Peripheral NeP | n = 65 | 12-Item Multiple Sclerosis Walking Scale | Reliability—Internal consistency and test-retest reliability; Validity—Hypothesis testing |
| Graham & Hughes⁶⁶ | Peripheral NeP | n = 100 | The Overall Neuropathy Limitations Scale | Reliability—Internal consistency, inter-rater, test-retest reliability; Validity—Content validity and Construct validity, Responsiveness |
| Harden et al.³⁶ | Complex and noncomplex regional pain syndrome | n = 155 | Complex regional pain syndrome severity score | Validity—Concurrent validity |
| Helme et al.²⁰ | Chronic Neuropathic Pain due to Postherpetic Neuralgia | n = 49 | McGill Pain Questionnaire | Validity—Concurrent validity |
| Jensen et al.²³ | Peripheral NeP | n = 133 | The Neuropathic Pain Scale | Validity—Responsiveness |
| Jensen et al.²⁴ | Diabetes-related foot pain | n = 159 | The Neuropathic Pain Scale | Validity—Responsiveness |
| Jensen et al.²⁵ | Carpal Tunnel Syndrome | n = 40 | Pain Quality Assessment Scale | Validity—Responsiveness |
| Jensen et al.²⁶ | Carpal Tunnel Syndrome | n = 100 | Pain Quality Assessment Scale | Reliability—Internal consistency; Validity—Construct validity |
| Kilmer et al.²⁷ | Hereditary motor and sensory NeP | n = 9 | Work stimulation tasks; Hand-held dynamometry; Brief Pain Inventory—Facial | Reliability—test-retest reliability; Validity—Construct validity |
| Lee et al.⁴⁹ | Typical and atypical facial pain due to Trigeminal Neuralgia | n = 156 | Physical Performance Measures | Reliability—test-retest reliability |
| Manor et al.⁵⁷ | Peripheral NeP | n = 20 | Quantitative sensory testing (thermal sensitivity) | Reliability—inter-rater reliability |
| Maser et al.³⁴ | Diabetic neuropathy | n = 100 | Screening of Activity Limitation and Safety Awareness Scale | Validity—Construct validity |
| Melchior & Velema⁷⁰ | Leprosy-related Neuropathic Pain | n = 25 | The INCAT Overall Disability Sum Score | Validity—Concurrent validity |
| Merkies & Schmitz⁶¹ | Guillain Barré Syndrome; Chronic idiopathic demyelinating polynoepathies | n = 20 | The Overall Disability Sum Score | Reliability—inter-rater and intrarater reliability; Validity—construct validity, Responsiveness |
| Merkies⁶⁵ | Neuropathic Pain | n = 113 | Inflammatory Sensory Score | Reliability—Internal consistency, inter-rater, intrarater reliability; Validity—Construct validity, Responsiveness |
| Merkies et al.⁶² | Neuropathic Pain | n = 113 | | |
| Mondelli et al.⁷⁶ | Ulnar Neuropathy at Elbow; Carpal Tunnel Syndrome | n = 292 | Ulnar neuropathy at the elbow Questionnaire | Reliability—Internal consistency and test-retest reliability; Validity—content validity and construct validity, Responsiveness |
| Murphy et al.⁵² | Charcot-Marie-Tooth disease | n = 34 | Charcot-Marie-Tooth disease neuropathy score-2 | Reliability—inter-rater and intrarater reliability |
| Novak et al.⁵⁵ | Peripheral Nerve injury | n = 124 | The Disabilities of Arm, Shoulder, and Hand Questionnaire | Reliability—Internal consistency; Validity—Construct validity |
| Reference               | Participant’s Characteristics | Outcome Measures Studied | Psychometric Properties Tested |
|-------------------------|------------------------------|--------------------------|--------------------------------|
| Novak et al.18          | Type 2 diabetes NeP          | n = 30                   | Foot Function Index (pain subscale) | Reliability—Internal consistency; Validity—Hypothesis testing |
|                         | Gender = 10 M, 20 F         |                          |                                | Validity—Inter-rater and test-retest reliability; Validity—Construct validity |
| Oerlemans et al.69      | Reflex Sympathetic Dystrophy Syndrome | n = 54                  | The Radvoud skills Questionnaire | Reliability—Internal consistency; Validity—Construct validity |
| Padua et al.47          | Charcot-Marie-Tooth disease | n = 211                  | Barthel Index; Deamallion Index | Validity—Construct validity |
|                         | Gender = 84 M, 127 F        |                          |                                | Validity—Construct validity |
| Padua et al.19          | Peripheral Nerve disease    | n = 392                  | Italian Neuropathic Pain       | Validity—Construct validity, Responsiveness |
|                         | Gender = 218 M, 174         |                          | Symptom Inventory              | Reliability—Test-retest reliability |
| Perez et al.68          | Complex regional pain syndrome-1 | n = 21                 | Walking stairs Questionnaire; Questionnaire rising and sitting down | Reliability—Internal consistency; Validity—Responsiveness |
|                         | Gender = 4 M, 17 F          |                          | Sheehan Disability Scale       | Validity—Hypothesis testing |
| Rejas et al.74          | Neuropathic Pain            | n = 603                  |                               | Reliability—Construct validity |
|                         | Gender = 211 M, 392 F       |                          |                               | |
| Schmader et al.45       | Herpes Zoster               | n = 165                  |                               | Reliability—Internal consistency and test-retest reliability; Validity—Construct validity |
|                         | Gender = 66 M, 99 F         |                          |                               | Validity—Inter-rater and intrarater reliability; Validity—Construct validity |
| Schreuders et al.36     | Charcot-Marie-Tooth disease | n = 45 Gender = 25 M, 20 F |                               | Validity—Responsiveness |
|                         | Gender = 25 M, 20 F         |                          |                               | |
| Sezgi'n et al.79 2006  | Idiopathic Carpal Tunnel Syndrome | n = 67                   | Turkish version of the Boston Questionnaire | Reliability—Construct validity |
|                         | Gender = 5 M, 62 F          |                          |                               | |
| Shy et al.70            | Charcot-Marie-Tooth disease | n = 60                   | Charcot-Marie-Tooth disease neuropathy score | Reliability—Inter-rater and test-retest reliability; Validity—Construct validity |
|                         | Gender = not mentioned       |                          | Charcot-Marie-Tooth disease | Validity—Construct validity |
| Shy et al.51            | Charcot-Marie-Tooth disease | n = 72                   | Neurpathy impairment score | Validity—Responsiveness |
|                         | Gender = 48 M, 24 F         |                          | The Utah Early neuropathy scale | |
| Singleton et al.42      | Diabetic peripheral NeP      | n = 129                  | Step Activity Monitor          | Reliability—Inter-rater reliability; Validity—Criterion validity, Responsiveness |
|                         | Gender = not mentioned       |                          |                               | Validity—Hypothesis testing |
| Smith et al.72          | Diabetic peripheral NeP      | n = 57                   |                               | Reliability—Inter-rater and intrarater reliability |
|                         | Gender = 57 M, 0 F          |                          |                               | |
| Solari et al.64         | Charcot-Marie-Tooth disease | n = 40                   | The Overall Neuropathy Limitations Scale; 10 m walk; 9 hole peg test | Reliability—Internal consistency; Validity—Content validity |
|                         | Gender = 21 M, 19 F         |                          | Screening of Activity Limitation and Safety Awareness Scale | Reliability—Test-retest reliability; Validity—Construct validity |
| Group SCS 2007          | Leprosy- and Diabetes-related NeP | n = 568                 | The Diabetes symptom checklist Type-2 | Validity—Construct validity, Validity—Construct validity and Criterion validity |
| Valk et al.17           | Type I and II diabetes NeP  | n = 78                   | Step Activity Monitor (4 minute walking test) | |
|                         | Gender = 34 M, 35 F         |                          |                               | |
| van Schie et al.73      | Diabetic peripheral neuropathy | n = 24                  |                               | Reliability—Internal Consistency; Validity—Construct validity |
|                         | Gender = 17 M, 7 F          |                          |                               | Validity—Construct Validity |
| VanSwearingen & Brach78 | Facial paralysis             | n = 46                   | Facial Disability Index        | Reliability—Internal Consistency, test-retest reliability; Validity—Construct validity |
|                         | Gender = 16 M, 30 F         |                          |                               | |
| Videler et al.79        | Hereditary motor and sensory type 1a neuropathy | n = 49                 | Sollerman Hand function test; | Reliability—Internal Consistency, test-retest reliability; Validity—Construct validity |
|                         | Gender = 21 M, 28 F         |                          | Functional dexterity test      | |
| Villoria et al.38       | Chronic Neuropathic Pain    | n = 548                  | Spanish Neuropathic Pain       | |
|                         | Gender = 209 M, 339 F       |                          | Symptom Inventory              | |
| Zelman et al.15         | Diabetic Peripheral NeP      | n = 255                  | Brief Pain Inventory-Diabetic Peripheral Neuropathy scale | Reliability—Internal Consistency; Validity—Construct validity, Discriminative and Criterion validity |
|                         | Gender = 114 M, 131 F       |                          |                               | Validity—Construct validity and Construct validity |
| Zimmerman et al.56      | Ulnar nerve injury          | n = 48                   | The Disabilities of the Arm, Shoulder, and Hand Questionnaire; Levine-Katz Questionnaire | |
|                         | Gender = not mentioned       |                          |                               | |
of health status Measurement Instruments (COSMIN)\textsuperscript{6} checklist; developed by an international group of experts, is unique and preferred because it allows for individual assessment of each psychometric domain within a study.

The COSMIN checklist\textsuperscript{14} (Table 3) consists of “A to J” 10 boxes: (Internal consistency—Box A; Reliability—Box B; Measurement error—Box C; Content validity—Box D; Structural validity—Box E; Hypotheses testing—Box F; Cross-cultural validity—Box G; Criterion validity—Box H; Responsiveness—Box I; Interpretability—Box J), with 5 to 18 items concerning methodological standards for how each measurement property should be assessed. According to COSMIN guidelines, the methodological quality of a study is considered adequate if all items in a box (A to J) were considered adequate. For this, each item was scored on a 4-point rating scale (ie, “poor,” “fair,” “good,” or “excellent”). The primary investigator (PM) independently scored all articles and the results were discussed and consensus obtained with each relevant team member. Methodological quality was determined using the “lowest rating score”\textsuperscript{6} achieved by any item for the representative psychometric property. Therefore, if one criterion for any property scored “poor”, the methodological quality for that particular property was rated as “poor” overall, irrespective of the scores that other criteria achieved. Disagreements regarding COSMIN scoring were resolved by discussion between reviewers. Reviewers were not blinded to the journal affiliation or authors of the included articles.

### RESULTS

Figure 1 illustrates the study selection process. The search resulted in 10,913 articles. After accounting for duplicate removal, title screening, and abstract screening, 80 articles were identified and retrieved as potentially eligible for the review. While checking the eligibility of full-text articles, a further 16 articles were excluded from the review as 2 articles were editorial papers; 2 were commentary papers; 5 articles were based on cancer pain; 3 articles were PhD publications; and for the remaining 4, full-text articles were not available. Thus, a total of 64 articles satisfied our eligibility criteria and were included in this review.

#### Characteristics of Included Studies

In total, 64 studies reporting 61 different outcome measures were identified. The included studies evaluated the psychometric properties of pain outcome domains ($n=24$) and physical function outcome domains ($n=37$) (Table 2). For the 24 pain intensity outcome measures, 15 (63\%) measures were patient-reported/self-reported measures, and the other 9 (37\%) were the therapist/clinician completed measures. For the 37 physical function outcome measures, 17 (46\%) measures were patient-reported/self-reported measures, that is, symptomatic assessment (subjective), 9 (24\%) measures were performance-based measures, and the rest of the 11 (30\%) measures were therapist completed measures, that is, symptoms and signs (subjective and objective testing). The synthesis of results per/outcome measure, their published psychometric properties, and quality assessment scores for studies are detailed in Tables S1 & S2. Data on the characteristics of the study population and sample population were extracted on the interpretability and generalizability boxes provided by the COSMIN checklist. Information regarding the sample size and gender distribution is reported in Table 2.

#### Pain intensity Outcome Measures

Pain domain outcomes (Tables 2 and S1) included the following: Brief Pain Inventory Scale for Diabetic Peripheral Neuropathy;\textsuperscript{15} Complex Regional Pain Syndrome Severity Score;\textsuperscript{16} Diabetes Symptom Checklist Type-2;\textsuperscript{17} Foot Function Index (pain subscale);\textsuperscript{18} Italian Neuropathic Pain Symptom Inventory;\textsuperscript{19} McGill Pain Questionnaire;\textsuperscript{20} modified Toronto Clinical Neuropathy
Score;²¹ Neuropathic Pain Scale;²²–²⁴ Neuropathic Pain Sensory Inventory;²⁵,²⁶ 0–10 Numerical Rating Scale;²⁷ Neuropathy Total Symptom Score-6;²⁸ 0–10 point Pain Intensity—Numerical Rating Scale;²⁹ Pain Quality Assessment Scale;³⁰,³¹ Portuguese version of the Neuropathic Pain Symptoms Inventory;³² Quantitative Sensory Testing (hot and cold pain threshold);³³–³⁵ Sensory evaluation with Semmes-Weinstein Monofilaments;³⁶ Short-form McGill Pain Questionnaire-2;³⁷ Spanish Neuropathic Pain Symptom Inventory;³⁸ Toronto Clinical Scoring System;³⁹ Total Neuropathy Score;⁴⁰ Trauma Related Neuronal Dysfunction Symptoms Inventory;⁴¹ Utah Early Neuropathy Scale;⁴² Visual analog scale;⁴³ and Zoster Brief Pain Inventory.⁴⁴,⁴⁵

Physical Functioning Outcome Measures

The range of physical functioning outcome measures was equally extensive and included (Tables 2 and S2): Alderson–McGall Hand Function questionnaire;⁴⁶
Barthel Index;47 Berg Balance Measure;48 Brief Pain Inventory Facial;49 Charcot–Marie–Tooth disease Neuropathy score;50,51 Charcot–Marie–Tooth disease Neuropathy Score-2,52 Disabilities of Arm, Shoulder, and Hand Questionnaire;53–56 Deambulation Index;47 Dellon-modified Moberg pick-up test;57 Facial Disability Index;58 Functional Dexterity test;59 Human Activity Profile;60 INCAT The Overall Disability Sum Score;61 Inflammatory neuropathy Sensory Score;62 Levine-Katz Questionnaire;63 Michigan Hand Outcome Questionnaire;53 modified Neuropathy Disability Score;63 10-m walking test;48,64 Nine-Hole Peg test;64 Neuropathy Impairment Score;65 Overall Disability Sum Score;66 Overall Neuropathy Limitations Scale;64,66 Patient Evaluation Measure;53 Physical Performance Measures (6 minute walk test, Timed up and go test);67 Questionnaire Rising and Sitting down;68 Radboud skills Questionnaire;69 short form Screening of Activity Limitation and Safety Awareness Scale;70,71 Step Activity Monitor;72 Step Activity Monitor (4-min walk test);73 Sheehan Disability Scale;74 Sollerman Hand function test;75 Turkish version of the Boston Questionnaire;75 Ulnar Neuropathy at the Elbow Questionnaire;76 12-Item Multiple Sclerosis Walking Scale;77 Walking Stairs Questionnaire;68 Work stimulation tasks (knob turn, Linear motion, and Lever arm);78 and Zoster Impact Questionnaire.45

Methodological Quality of Studies Evaluating Psychometric Properties of Pain Intensity and Physical Functioning Outcome Measures

Reliability. The majority of the instruments included in our review were not tested for all psychometric properties listed on COSMIN checklist. Forty-four of the 64 studies (68%) assessed various forms of reliability (Internal consistency, inter-rater reliability, intrarater reliability, test–retest reliability, and measurement error) and showed a mixed methodological quality of evidence (excellent/good/fair/poor), when evaluated on COSMIN (Tables S1 & S2). The key results for reliability showed that the BPI-DPN and the SF-MPQ2 have excellent (α > 0.90) internal consistency. The mTCNS has good internal consistency (α = 0.81 to 0.90), inter-rater reliability, and intrarater reliability (ICC or K = 0.81 to 0.90). The hot and cold pain thresholds on the QST have good inter-rater and test–retest reliability (ICC or K = 0.81-0.90). The Spanish NPSI has excellent internal consistency (α > 0.90) with good test–retest reliability (ICC or K = 0.81 to 0.90).

Measurement error was the least reported form of reliability, and the TRNDSI had good test–retest reliability (ICC or K = 0.81 to 0.90) and measurement error (see Table S1). These measures with excellent and good psychometric properties scores also scored good/excellent on the COSMIN checklist (according to COSMIN criteria stated in Table S3).

Validity. Validity was the more frequently tested psychometric property, in 49 of 64 studies (76%), there was face/content validity, structural validity, construct validity, criterion/concurrent validity, convergent validity, discriminative validity, hypothesis testing, and responsiveness. Similar to the findings for reliability, mixed methodological quality evidence (excellent/good/fair/poor) was found when evaluated on COSMIN (Tables S1 and S2). The key results for validity showed that the NPSI, the SALSA, and the UNEQ have excellent content validity as there were no concerns raised by the patients or experts regarding the wording of questionnaires, and thus, no further modifications were advised. The UENS has the best criterion validity followed by the HAP and the mNDS. Approximately one-third of the studies (18/49, 36%) evaluated responsiveness form of validity. The NPSI has excellent responsiveness followed by the 0 to 10 PI-NRS and the ODSS. Also, the studies showing these evidences were of excellent/good methodological quality on the COSMIN checklist (as according to COSMIN criteria stated in Table S3).

DISCUSSION

To our knowledge, this is the first systematic review to evaluate the evidence for the psychometric properties of pain and physical functional outcome measures used in assessment in NeP conditions and to identify the methodological quality of the studies investigating the psychometric properties of various outcome measures. A total of 61 different outcome measures were identified related to the domains of pain and physical functioning. In this systematic review, while most of the studies have shown good/excellent evidence of reliability and validity of the used scales, only few are considered “excellent to good” in terms of their methodological quality. Our review identified acceptable reliability and validity (for a few key properties) for the mTCNS, the TRNDSI, the 0-10 PI NPS, the QST, the SALSA, the Spanish NPSI, the ODSS, the SF-MPQL, the UNEQ, the UENS, the HAP, the mNDS, the NDS, and the BPI-DPN.
The available studies investigating the psychometric property of reliability were rated in varying methodological quality from “poor” to “excellent” on the COSMIN checklist. However, the majority of studies showed similar methodological shortcomings. In this review, smaller sample sizes were found to be associated with the majority of inconsistent results. According to COSMIN guidelines, a sample size of ≥ 100 is considered to be an adequate/excellent sample size, given the need for precision in the overall estimates; these estimates are based on the power 0.80. A sample size of 50 provides a 0.70 power (level of significance being 0.05), while 100 has a power of 0.94.

In the current systematic review, many outcome measures seem promising for different domains of reliability and validity (according to COSMIN criteria stated in Table S3), as the FFI, the NTSS-6, the AMHFQ, the DASH, the HAP, the ISS, the MHQ, the PEM, the SDS, the TBQ, the UNEQ, and the Walk-12 scales have “moderate” (α ≥ 0.71 to 0.80) to “excellent” (α ≥ 0.90) published grades for internal consistency. However, when the methodological quality of the studies was evaluated on COSMIN, these were graded of “poor/fair” quality because of the small sample size. These findings are consistent with those of a recent systematic review on outcome measures in neck pain, where smaller sample sizes frequently led to poorer results. This current review recommends that future research on a larger sample size (n ≥ 100, as recommended by COSMIN) is needed to improve the quality of research on these measures.

Validity was the most frequently evaluated psychometric property in both pain and physical functioning outcome domains. The majority of these studies demonstrated unsatisfactory (poor/fair scores) results on COSMIN. The main reasons for this were inconsistencies in the following areas: smaller sample sizes; hypotheses were not formulated; and expected direction/magnitude of correlations was not stated in advance. Other common findings were a lack of information about reporting of missing items, and measures adopted to handle missing data. Although these 2 items did not contribute to the overall “poor” grading on the COSMIN, it is expected that studies of “good” methodological quality should report this construct, as a high number of missing items can introduce bias.

A further interesting finding of this review was that responsiveness was the least frequently studied psychometric property for the included pain and physical functioning outcome measures. There were a total of 18 studies which published the findings on responsiveness and only 3 scales—the NPS, the 0 to 10 PI-NRS, and the ODSS proved satisfactory methodological quality on COSMIN. The remaining measures were graded “fair to poor”, and all the above-stated shortcomings (small sample size, un-reporting of missing items, vagueness about how the missing data were handled, not well-formulated hypothesis, etc.) equally contributed to the inconsistent results for the studies reporting on this property.

In the current systematic review, there were few measures identified which had promising psychometric properties for key variables; the mTCNS (good internal consistency, inter-rater and intrarater reliability, and criterion validity); the TRNDSI, and the ZBPI (good test–retest reliability); the NPS (excellent face/content validity); the 0 to 10 PI-NRS (good responsiveness); the QST—pain threshold (good intrarater and test–retest reliability); the NPS (excellent responsiveness); and the SALSA (excellent internal consistency and content validity), and were supported by a “excellent to good” methodological quality on the COSMIN checklist. The future use of these measures can be recommended based on their proven psychometric properties; however, it is imperative that other remaining psychometric properties of these outcome measures should also be established.

We also identified a list of instruments which showed their best methodological quality for few psychometric properties on COSMIN, but at the same time good methodological quality evidence was lacking for other properties: the TCSS (good construct validity, but poor inter- and intrarater reliability); the Short-form MPQ-2 (excellent internal consistency, but fair construct validity and responsiveness); the HAP (good criterion validity, with poor internal consistency and responsiveness, and fair hypothesis testing); the ODSS (good responsiveness but fair inter-rater and intrarater reliability and construct validity); the UNEQ (excellent content validity, fair test–retest reliability, and poor internal consistency, construct validity, and responsiveness); the BPI-DPN (good construct validity, fair test–retest reliability, and poor internal consistency and discriminative validity, fair construct validity and poor criterion validity). As study methodology may influence results for psychometric properties, it is recommended that further evaluation of these psychometric properties with studies of improved methodological quality should be carried out.
LIMITATIONS

Firstly, it is acknowledged that “Neuropathic Pain conditions” is an umbrella term which covers a range of different conditions such as diabetic neuropathy, trigeminal neuralgia, and postherpetic neuralgia. For the search strategy, MESH terms/keywords indexed for neuropathy, neuralgia, and neurodynia were used to be as inclusive as possible. It is acknowledged that each condition could have been separately searched and that such an approach may have lessened the chances of missing studies.

Secondly, psychometric properties such as reliability and validity, including responsiveness, are subclassified into various forms such as internal consistency, inter-rater/re-test reliability, content validity, minimal important difference, and standard error of measurement, etc. For the current search strategy, keywords in 3 broader areas (reliability, validity, and/or responsiveness) were used rather than individual subclassified keywords. However, as these broader terms are the most commonly used to denote the various forms of psychometric properties, it is anticipated that the majority of studies would have been selected.

Lastly, for this systematic review, multidisciplinary, international consensus-based methodological quality reporting guidelines, and COSMIN were followed for rating the quality of included studies of psychometric properties. The COSMIN checklist has well developed data extraction forms with detailed instructions for completion. The 4-point rating scale classifies each assessment of a measurement property as “excellent, good, fair, or poor”, based on the scores of the items in the corresponding COSMIN box. The methodological quality of a study is considered adequate if all items in a box (A to J) are considered adequate. However, frequently not all items in a box are scored adequate, and it is not feasible to provide overall definitive grade for each psychometric property; thus, no decisions can be drawn for the methodological quality of the studies-based purely on COSMIN findings.

CONCLUSION

In this review, we evaluated the evidence for psychometric properties of 61 unique outcome measures identified to assess pain and physical functioning outcome domains in trials of NeP conditions. We have presented extensive data, which demonstrate the psychometric properties of these available outcome measures, and recommend the use of the mTCNS, the TRNDSI, the ZBPI, the NPSI, the 0 to 10 PI-NRS, the QST—pain threshold, and the NPS to detect changes in pain intensity and physical functions. We found that important information regarding the methodological quality of the majority of studies demonstrating these psychometric properties is lacking or is of poor quality. As NeP is a multidisabling condition with significant associated morbidity, usage of quality-evidenced pain and physical functional measures is a key recommendation for future research in NeP intervention studies. It appears that despite representing these measures in many studies of NeP, the methodological quality for most of the measures is not strong enough to recommend their use based on their psychometric properties. Thus, good quality future research is required to further investigate the psychometric properties of identified outcome measures used for clinical and research purposes.

ACKNOWLEDGEMENTS

All authors declare that there exists no conflict of interests associated with the current study. The authors thank and acknowledge Dr Ramakrishnan Mani, Lecturer, Centre for Health Activity and Rehabilitation Research, School of Physiotherapy, University of Otago for his suggestions, invaluable constant assistance and helping with the constructive feedback on drafts of the manuscript. The findings of the study have been presented as a poster in the 8th Congress of the European Federation of IASP® Chapters (EFIC 2013) in Florence, Italy.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1 Summary of identified pain Intensity outcome measures with their published psychometric properties and COSMIN grading.

Table S2 Summary of identified physical functioning outcome measures with their published psychometric properties and COSMIN grading.

Table S3 Definition of domains, measurement properties, aspects of measurement properties and accepted statistical analyses by COSMIN.

REFERENCES

1. Treede RD, Jensen TS, Campbell JN, et al. Neuropathic pain: redefinition and a grading system for clinical and research purposes. Neurology. 2008;70:1630–1635.
2. Turk DC, Dworkin RH, Allen RR, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. Pain. 2003;106:337–345.

3. Cruccu G, Sommer C, Anand P, et al. EFNS guidelines on neuropathic pain assessment: revised 2009. Eur J Neurol. 2010;17:1010–1018.

4. Haanpaa M, Attal N, Backonja M, et al. NeuPSIG guidelines on neuropathic pain assessment. Pain. 2011;152:14–27.

5. Roy JS, Desmeules F, MacDermid JC. Psychometric properties of presenteeism scales for musculoskeletal disorders: a systematic review. J Rehabil Med. 2011;43:23–31.

6. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63:737–745.

7. Stinson JN, Kannavagh T, Yamada J, Gill N, Stevens B. Systematic review of the psychometric properties, interpretability and feasibility of self-report pain intensity measures for use in clinical trials in children and adolescents. Pain. 2006;125:143–157.

8. Squires JE, Estabrooks CA, O’Rourke HM, Gustavsson P, Newburn-Cook CV, Wallin L. A systematic review of the psychometric properties of self-report research utilization measures used in healthcare. Implement Sci. 2011;6:83.

9. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res. 2012;21:651–657.

10. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess. 1998;2:i–iv, 1–74.

11. Gore M, Brandenburg NA, Hoffman DL, Tai KS, Stacey B. Burden of illness in painful diabetic peripheral neuropathy: the patients’ perspectives. J Pain. 2006;7:892–900.

12. Mehta P, Claydon L, Hendrick P, Winser S, Baxter GD. Outcome Measures in Randomized Controlled Trials of Neuropathic Pain Conditions: a Systematic Review of Systematic Reviews and Recommendations for Practice. Clin J Pain. 2015;31:169–176.

13. Clinical Resource Efficiency Support Team (CREST). Guidelines on the management of neuropathic pain. 2008. http://www.thblack.com/links/RSD/CRESTManagementNeuropathicPainGuidelines.pdf (accessed September, 2012)

14. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60:34–42.

15. Zelman DC, Gore M, Dukes E, Tai KS, Brandenburg N. Validation of a modified version of the brief pain inventory for painful diabetic peripheral neuropathy. J Pain Symptom Manage. 2005;29:401–410.

16. Harden RN, Bruehl S, Perez RS, et al. Development of a severity score for CRPS. Pain. 2010;151:870–876.

17. Valk GD, Groothuis PA, van Eijik JT, Bouter LM, Bertelsmann FW. Methods for assessing diabetic polyneuropathy: validity and reproducibility of the measurement of sensory symptom severity and nerve function tests. Diabetes Res Clin Pract. 2000;47:87–95.

18. Novak P, Burger H, Marincek C, Meh D. Influence of foot pain on walking ability of diabetic patients. J Rehabil Med. 2004;36:249–252.

19. Padua L, Brian C, Jann S, et al. Validation of the Italian version of the Neuropathic Pain Symptom Inventory in peripheral nervous system diseases. Neurol Sci. 2009;30:99–106.

20. Helme RD, Katz B, Gibson S, Corran T. Can psychometric tools be used to analyse pain in a geriatric population? Clin Exp Neurol. 1989;26:113–117.

21. Bril V, Tomioka S, Buchanan RA, Perkins BA; m TSG. Reliability and validity of the modified Toronto Clinical Neuropathy Score in diabetic sensorimotor polyneuropathy. Diabet Med. 2009;26:240–246.

22. Galer BS, Jensen MP. Development and preliminary validation of a pain measure specific to neuropathic pain: the Neuropathic Pain Scale. Neurology. 1997;48:332–338.

23. Jensen MP, Dworkin RH, Gammaitoni AR, Olaluye DO, Oleka N, Galer BS. Assessment of pain quality in chronic neuropathic and nociceptive pain clinical trials with the Neuropathic Pain Scale. J Pain. 2005;6:98–106.

24. Jensen MP, Friedman M, Bonzo D, Richards P. The validity of the neuropathic pain scale for assessing diabetic neuropathic pain in a clinical trial. Clin J Pain. 2006;22:97–103.

25. Bouhassira D, Attal N, Ferneman J, et al. Development and validation of the Neuropathic Pain Symptom Inventory. Pain. 2004;108:248–257.

26. Crawford B, Bouhassira D, Wong A, Dukes E. Conceptual adequacy of the neuropathic pain symptom inventory in six countries. Health Qual Life Outcomes. 2008;6:62.

27. Farrar JT, Pritchett YL, Robinson M, Prakash A, Chappell A. The clinical importance of changes in the 0 to 10 numeric rating scale for worst, least, and average pain intensity: analyses of data from clinical trials of duloxetine in pain disorders. J Pain. 2010;11:109–118.

28. Basyer EJ 3rd, Price KL, Bril V, Group MS. Development and validity testing of the neuropathy total symptom score-6: questionnaire for the study of sensory symptoms of diabetic peripheral neuropathy. Clin Ther. 2005;27:1278–1294.

29. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain. 2001;94:149–158.

30. Jensen MP, Gammaitoni AR, Olaluye DO, Oleka N, Nalamachu SR, Galer BS. The pain quality assessment scale: assessment of pain quality in carpal tunnel syndrome. J Pain. 2006;7:823–832.

31. Jensen MP, Gould EM, Victor TW, Gammaitoni AR, White RE, Galer BS. The relationship of changes in pain quality to pain interference and sleep quality. J Pain. 2010;11:782–788.

32. de Andrade DC, Ferreira KA, Nishimura CM, et al. Psychometric validation of the Portuguese version of the
Neuropathic Pain Symptoms Inventory. *Health Qual Life Outcomes*. 2011;9:107.

33. Geber C, Klein T, Azad S, et al. Test-retest and interobserver reliability of quantitative sensory testing according to the protocol of the German Research Network on Neuropathic Pain (DFNS): a multi-centre study. *Pain*. 2011;152:548–556.

34. Maser RE, Nielsen VK, Bass EB, et al. Measuring diabetic neuropathy. Assessment and comparison of clinical examination and quantitative sensory testing. *Diabetes Care*. 1989;12:270–275.

35. Felix ER, Widerstrom-Noga EG. Reliability and validity of quantitative sensory testing in persons with spinal cord injury and neuropathic pain. *J Rehabil Res Dev*. 2009;46:69–83.

36. Schreuders TA, Selles RW, van Ginneken BT, Janssen WG, Stam HJ. Sensory evaluation of the hands in patients with Charcot-Marie-Tooth disease using Semmes-Weinstein monofilaments. *J Hand Ther*. 2008;21:28–34; quiz 35.

37. Dworkin RH, Turk DC, Revicki DA, et al. Development and initial validation of an expanded and revised version of the Short-form McGill Pain Questionnaire (SF-MPQ-2). *Pain*. 2009;144:35–42.

38. Villoria J, Rodriguez M, Berro MJ, Stern A, Sanchez-Magro I. Psychometric validation of the neuropathic pain symptom inventory for its use in Spanish. *J Pain Symptom Manage*. 2011;42:134–146.

39. Bril V, Perkins BA. Validation of the Toronto Clinical Scoring System for diabetic polyneuropathy. *Diabetes Care*. 2002;25:2048–2052.

40. Cornblath DR, Chaudhry V, Carter K, et al. Total neuropathy score: validation and reliability study. *Neurology*. 1999;53:1660–1664.

41. Collins S, van Hilten JJ, Marinus J, Zuurmond WW, de Lange JJ, Perez RS. Development of a symptoms questionnaire for complex regional pain syndrome and potentially related illnesses: the Trauma Related Neuronal dysfunction inventory. *J Pain Symptom Manage*. 2008;9:1114–1120.

42. Singleton JR, Bixby B, Russell JW, et al. The Utah Early Neuropathy Scale: a sensitive clinical scale for early sensory predominant neuropathy. *J Peripher Nerv Syst*. 2008;13:218–227.

43. Davidoff G, Morey K, Anmann M, Stamps J. Pain measurement in reflex sympathetic dystrophy syndrome. *Pain*. 1988;32:27–34.

44. Coplan PM, Schmader K, Nikas A, et al. Development of a measure of the burden of pain due to herpetic zoster and postherpetic neuralgia for prevention trials: adaptation of the brief pain inventory. *J Pain*. 2004;5:344–356.

45. Schmader KE, Sloane R, Pieper C, et al. The impact of acute herpes zoster pain and discomfort on functional status and quality of life in older adults. *Clin J Pain*. 2007;23:490–496.

46. Alderson M, McGill D. The Alderson-McGill hand function questionnaire for patients with Carpal Tunnel syndrome: a pilot evaluation of a future outcome measure. *J Hand Ther*. 1999;12:313–322.

47. Padua L, Aprile I, Cavallaro T, et al. Relationship between clinical examination, quality of life, disability and depression in CMT patients: Italian multicenter study. *Neuroul Sci*. 2008;29:157–162.

48. Erdmann PG, van Meeteren NL, Kalmijn S, Wokke JH, Holders PJ, van den Berg LH. Functional health status of patients with chronic inflammatory neuropathies. *J Peripher Nerv Syst*. 2005;10:181–189.

49. Lee JY, Chen HI, Urban C, et al. Development of and psychometric testing for the Brief Pain Inventory-Facial in patients with facial pain syndromes. *J Neurosurg*. 2010;113:516–523.

50. Shy ME, Blake J, Krajewski K, et al. Reliability and validity of the CMT neuropathy score as a measure of disability. *Neurology*. 2005;64:1209–1214.

51. Shy ME, Chen L, Swan ER, et al. Neuropathy progression in Charcot-Marie-Tooth disease type 1A. *Neurology*. 2008;70:378–383.

52. Murphy SM, Herrmann DN, McDermott MP, et al. Reliability of the CMT neuropathy score (second version) in Charcot-Marie-Tooth disease. *J Peripher Nerv Syst*. 2011;16:191–198.

53. Dias JJ, Rajan RA, Thompson JR. Which questionnaire is best? The reliability, validity and ease of use of the Patient Evaluation Measure, the Disabilities of the Arm, Shoulder and Hand and the Michigan Hand Outcome Measure. *J Hand Surg Eur*. 2008;33:9–17.

54. Eklund E, Svensson E, Hager-Ross C. Hand function and disability of the arm, shoulder and hand in Charcot-Marie-Tooth disease. *Disabil Rehabil*. 2009;31:1955–1962.

55. Novak CB, Anastakis DJ, Beaton DE, Mackinnon SE, Katz J. Relationships among pain disability, pain intensity, illness intrusiveness, and upper extremity disability in patients with traumatic peripheral nerve injury. *J Hand Surg Am*. 2010;35:1633–1639.

56. Zimmerman NB, Kaye MB, Wilgis EF, Zimmerman RM, Dubin NH. Are standardized patient self-reporting instruments applicable to the evaluation of ulnar neuropathy at the elbow? *J Shoulder Elbow Surg*. 2009;18:463–468.

57. Amirjani N, Ashworth NL, Olson JL, Morhart M, Chan KM. Discriminative validity and test-retest reliability of the Dellon-modified Moberg pick-up test in carpal tunnel syndrome patients. *J Peripher Nerv Syst*. 2011;16:51–58.

58. VanSwearingen JM, Brach JS. The Facial Disability Index: reliability and validity of a disability assessment instrument for disorders of the facial neuromuscular system. *Phys Ther*. 1996;76:1288–1298; discussion 1298-1300.

59. Videler AJ, Beelen A, van Schaik IN, de Visser M, Nollet F. Manual dexterity in hereditary motor and sensory neuropathy type 1a: severity of limitations and feasibility and reliability of two assessment instruments. *J Rehabil Med*. 2008;40:132–136.

60. Farrell MJ, Gibson SJ, Helme RD. Measuring the activity of older people with chronic pain. *Clin J Pain*. 1996;12:6–12.
61. Merkies IS, Schmitz PI. Getting closer to patients: the INCAT Overall Disability Sum Score relates better to patients’ own clinical judgement in immune-mediated polyneuropathies. J Neurol Neurosurg Psychiatry. 2006;77:970–972.

62. Merkies IS, Schmitz PI, van der Meche FG, van Doorn PA. Psychometric evaluation of a new sensory scale in immune-mediated polyneuropathies. Inflammatory Neuropathy Cause and Treatment (INCAT) Group. Neurology. 2000;54:943–949.

63. Asad A, Hameed MA, Khan UA, Ahmed N, Butt MU. Reliability of the neurological scores for assessment of sensorimotor neuropathy in type 2 diabetics. J Pak Med Assoc. 2010;60:166–170.

64. Solari A, Laura M, Salsano E, Radice D, Pareyson D, Group C-TS. Reliability of clinical outcome measures in Charcot-Marie-Tooth disease. Neuromuscul Disord. 2008;18:19–26.

65. Merkies ISJ. Clinimetric evaluation of a new overall disability scale in immune mediated polyneuropathies. J Neurol Neurosurg Psychiatry. 2002;72:596–601.

66. Graham RC, Hughes RA. A modified peripheral neuropathy scale: the Overall Neuropathy Limitations Scale. J Neurol Neurosurg Psychiatry. 2006;77:973–976.

67. Manor B, Doherty A, Li L. The reliability of physical performance measures in peripheral neuropathy. Gait Posture. 2008;28:343–346.

68. Perez R, Roorda LD, Zuurmond WWA, Bannink I, Vranken JH, de Lange JJ. Measuring perceived activity limitations in lower extremity Complex Regional Pain Syndrome type 1 (CRPS I): test–retest reliability of two questionnaires. Clin Rehabil. 2002;16:454–460.

69. Oerlemans HM, Cup EH, DeBoo T, Goris RJ, Oostendorp RA. The Radboud skills questionnaire: construction and reliability in patients with reflex sympathetic dystrophy of one upper extremity. Disabil Rehabil. 2000;22:233–245.

70. Melchior H, Velem J. A comparison of the Screening Activity Limitation and Safety Awareness (SALSA) scale to objective hand function assessments. Disabil Rehabil. 2011;33:2044–2052.

71. Group SCS, Ebenso J, Fuzikawa P, et al. The development of a short questionnaire for screening of activity limitation and safety awareness (SALSA) in clients affected by leprosy or diabetes. Disabil Rehabil. 2007;29:689–700.

72. Smith DG, Domholdt E, Coleman KL, Del Aguila MA, Boone D. Ambulatory activity in men with diabetes: relationship between self-reported and real-world performance-based measures. J Rehabil Res Dev. 2004;41:571.

73. van Schie CH, Noordhof EL, Busch-Westbroek TE, Beelen A, Nollet F. Assessment of physical activity in people with diabetes and peripheral neuropathy. Diabetes Res Clin Pract. 2011;92:e9–e11.

74. Rejas J, Pardo A, Ruiz MA. Standard error of measurement as a valid alternative to minimally important difference for evaluating the magnitude of changes in patient-reported outcomes measures. J Clin Epidemiol. 2008;61:350–356.

75. Sezgi˘n M, İncel NA, Sevi˘m S, Çamdevi’ren H, As İ, Erdo˘Gan C. Assessment of symptom severity and functional status in patients with carpal tunnel syndrome: reliability and validity of the Turkish version of the Boston questionnaire. Disabil Rehabil. 2006;28:1281–1286.

76. Mondelli M, Padua L, Giannini F, Bibbo G, Aprili I, Rossi S. A self-administered questionnaire of ulnar neuropathy at the elbow. Neuroil. Sci. 2006;27:402–411.

77. Graham RC, Hughes RA. Clinimetric properties of a walking scale in peripheral neuropathy. J Neurol Neurosurg Psychiatry. 2006;77:977–979.

78. Kilmer DD, Aitkens SG, Wright NC, McCrory MA. Simulated work performance tasks in persons with neuropathic and myopathic weakness. Arch Phys Med Rehabil. 2000;81:938–943.

79. Stevens J. Applied Multivariate Statistics for the Social Sciences. Hillsdale, NJ, USA: L. Erlbaum Associates Inc.; 2009.

80. Schellingerhout JM, Verhagen AP, Heymans MW, de Vet HC, Terwee CB. Measurement properties of disease-specific questionnaires in patients with neck pain: a systematic review. Qual Life Res. 2012;21:659–670.

81. Dworkin RH, Backonja M, Rowbotham MC, et al. Advances in neuropathic pain: diagnosis, mechanisms, and treatment recommendations. Arch Neurol. 2003;60:1524–1534.

82. Portney LG, Watkins MP. Foundations of Clinical Research Applications to Practice. 3rd revised United States ed. Upper Saddle River/US: Prentice Hall; 2007.