

RESEARCH REPORT

The Patient-Rated Wrist and Hand Evaluation: a systematic review of its validity and reliability

Julie Taylor *MHPrac with Distinction, PGDipHSc, BHSc (Physiotherapy)*
Senior Physiotherapist, Whangarei Physiotherapy Services, Whangarei,

Paula Kersten *PhD, MSc, BSc (Physiotherapy)*
Associate Professor Rehabilitation, Person Centred Research Centre, Auckland University of Technology, Auckland,

ABSTRACT

The Patient-Rated Wrist and Hand Evaluation is a region specific, patient reported outcome measure that aims to evaluate pain and disability of the wrist and hand. This review appraised the evidence for the validity and reliability of the Patient-Rated Wrist and Hand Evaluation as a measure of therapeutic outcomes in musculoskeletal conditions affecting the wrist and hand. Relevant studies were identified by a search of the literature and evaluated according to the Consensus standards for the selection of health status measurement instruments checklist. Five studies met the inclusion criteria for review. Four studies utilised Classical Test Theory and were in support of the reliability and validity of the Patient-Rated Wrist and Hand Evaluation. However, all of these studies were of fair to poor methodological quality. A fifth study, which utilised Rasch analysis, was of good methodological quality and supported the validity and reliability of the Patient-Rated Wrist and Hand Evaluation as a three subscale instrument. In conclusion, there is some good evidence for the validity and reliability of the Patient-Rated Wrist and Hand Evaluation. Further work needs to be done to enable clinicians to rescore the tool into three subscales and to examine its cultural validity.

Taylor J, Kersten P (2014) The Patient-Rated Wrist and Hand Evaluation: a systematic review of its validity and reliability New Zealand Journal of Physiotherapy 42(2): 139-145.

Key words Patient-Rated Wrist and Hand Evaluation, Psychometrics, Rasch analysis, Hand injuries, Patient outcome assessment

INTRODUCTION

The routine evaluation of outcomes in hand therapy is important due to an increasing demand to justify the quality and cost effectiveness of health services, to direct treatment plans and to communicate therapeutic outcomes with patients and colleagues (Accident Compensation Corporation 2009, Amadio 2001, Jette 2009, Physiotherapy Board of New Zealand 2009). In hand therapy, pain and loss of function are key considerations when implementing an intervention, such as splinting, mobilisation, or exercise. The Patient-Rated Wrist and Hand Evaluation (PRWHE) is a commonly used outcome measure in hand therapy practice aiming to measure such outcomes. This review aimed to appraise the evidence for the validity and reliability of the PRWHE.

The PRWHE is a region specific outcome measure designed to evaluate pain and disability of the wrist and hand. It evolved from the Patient-Rated Wrist Evaluation (PRWE), which was originally developed and validated for conditions affecting the wrist. Both instruments consist of the same 15 items separated into two domains of pain (5 items) and function (10 items). However, the PRWHE makes reference to the wrist and/or hand as opposed to the wrist in isolation and includes two optional aesthetics questions. Function is further categorised as specific activities and usual activities (5 items each). An 11 point numerical scale (0-10) is utilised for each item. The scoring system is simple, whereby the functional scores are added and divided by two and then added to the pain scores, to give a total out of 100 (MacDermid and Tottenham 2004). Lower scores denote better function and less pain.

It is important that health related patient rated outcome (PRO) measures are developed from a strong conceptual basis, which

rationalises and clearly defines what and how it intends to measure (Holmbeck and Devine 2009, Rothman et al 2007). The conceptual basis and developmental process undertaken for the PRWE is thoroughly presented and discussed by MacDermid (1996). Although the author did not report the literature review strategy in detail, the review appropriately aimed to identify articles, which studied the physical requirements of various functional tasks and other patient reported outcome measures for pain and/or disability. MacDermid (1996) initially surveyed members of the International Wrist Investigators on current outcome measure use and opinions were sought regarding structure and content, promoting content validity and clinical utility of the measure. While MacDermid (1996) adequately promoted the International Wrist Investigators as an appropriate group to survey they did not reveal the size of the sample surveyed or the response rate. The surveyed members deemed pain, functional ability and patient satisfaction to be the most important subjective indicators of outcome.

A PRO measure may be useful in determining the relative effectiveness of a particular intervention. There is an inherent degree of error with any type of measurement. Therefore, in order to correctly interpret the findings it is important to ascertain the validity, reliability and responsiveness of a measure, for a specific purpose, in a specific population (Horner and Larmer 2006). These terms are otherwise known as the psychometric properties of a measure and refer to the theoretical principles and rules as applied to measurement (Nunnally and Bernstein 1994). The PRWE has good psychometric properties in a number of conditions affecting the wrist (Hoang-Kim et al 2011). However, a standardised measure for a variety of conditions affecting both the wrist and hand would have greater clinical utility in a typical hand therapy

clinic as it would potentially negate the need to select specific instruments for different regions or conditions.

The aim of this review was to appraise the evidence for the validity and reliability of the PRWHE to inform measurement of therapeutic outcomes in musculoskeletal conditions affecting the wrist and hand.

METHODS

Literature Search Strategy

A literature search was conducted during September 2013 in Medline, Allied and Complementary Medicine Database (AMED), Cumulative Index to Nursing and Allied Health Literature (CINAHL) and the Cochrane Library with the aim of identifying relevant clinical articles for the review (Table 1). The search strategy included subject headings related to the target outcome measure and keywords related to the target study design (Table 2). Truncation was utilised where appropriate to allow for spelling variations. Booleans, AND and OR, were utilised to combine search terms. Scopus was also utilised to cross reference the reference lists of relevant articles forward and back.

Table 1: Electronic search for articles evaluating the psychometric properties of the PRWHE

Date	Database	Search Provider	Period	Results
19/09/2013	Medline	Ebsco	1998 onwards	44
19/09/2013	AMED	Ovid	1985 onwards	4
19/09/2013	CINAHL	Ebsco	1998 onwards	17
19/09/2013	Cochrane Library	Ovid		15

Allied and Complementary Medicine Database (AMED), Cumulative Index to Nursing and Allied Health Literature (CINAHL), Patient Rated Wrist and Hand Evaluation (PRWHE)

Table 2: Search strategy for articles evaluating the psychometric properties of the PRWHE

Subject Heading	Target Study Design
"patient rate* wrist evaluation"	valid*
OR "patient rate* wrist and hand evaluation"	OR reliab*
OR PRWE	AND OR responsive*
OR PRWHE	OR "classical test theory"
	OR "rasch analysis"

Inclusion Criteria

Studies designed to examine the validity and reliability of the PRWHE were included for review. Only studies in English were included.

Appraisal

Eligible studies were analysed according to the Consensus Standards for the Selection of Health Status Measurement Instruments (COSMIN) checklist. The COSMIN checklist is designed to assess the methodological quality of studies evaluating the psychometric properties of a health related outcome measure (Mokkink et al 2010a, Mokkink et al 2010b). It was developed using an international consensus process and has established reliability and validity (Mokkink et al 2010c). It is

increasingly used in outcome measurement research, including physiotherapy related reviews (Mijnarends et al 2013). The COSMIN tool is divided into sections for each psychometric property (such as internal consistency, reliability, and criterion validity); each ranked using a quality score. These scores are labelled as excellent, good, fair, and poor, and a detailed description of these labels for each psychometric property is given (Terwee et al 2012).

RESULTS

In total 80 citations were retrieved from the database search. After removing duplicates and full review of manuscripts, five articles were identified that met the inclusion criteria for review (Figure 1). A summary of their findings are presented in Table 3. A summary of the methodological quality of these studies according to the COSMIN checklist is presented in Table 4.

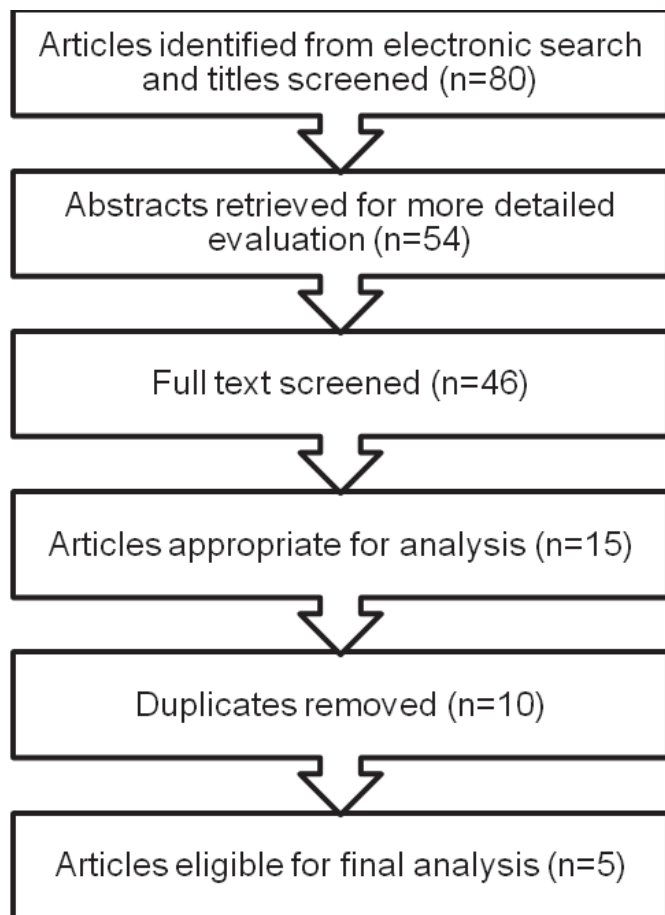
Validity of the PRWHE

Content validity refers to the degree to which the content of an instrument is an adequate reflection of the construct to be measured (Mokkink et al 2010b, Streiner and Norman 2008). While the PRWE was originally developed for wrist disorders,

MacDermid and Tottenham (2004) argue that from a functional perspective the wrist and hand are interrelated, and therefore, there is good face validity for its application in a population with hand injuries by changing the word "wrist" to "wrist/hand" and the name of the scale from PRWE to PRWHE. However, a formal evaluation of the face validity with clinicians or patients has not been conducted. The PRWHE includes pain as a measure of body structure and function limitation, and with respect to function, specific activities as a measure of ability and usual activities as a measure of participation. The inclusion of usual functional activities (personal care, household work, work and recreation), in addition to specific functional tasks, aims to encompass a wide range of tasks relevant to the individual. Additionally, the items are intended to cover a wide spectrum of each domain, for example, pain at rest designed to capture those with severe pain through to pain with a repeated wrist movement or lifting a heavy object designed to capture those with mild levels of pain (MacDermid 1996). When administering the PRWHE, difficulty may arise when a patient has not performed a certain activity, such as lifting a heavy object in the past week. In this case, patients are instructed to estimate the amount of pain or difficulty they would expect. If a patient has never performed an activity they can leave it blank. However, the content validity of the tool is affected by the potential of irrelevant items for some patients.

Criterion validity refers to how well a measure compares to a gold standard, either at the same time (concurrent validity) or one that will be available in the future (predictive validity)

Figure 1: Diagram of search selection process for relevant articles



(Streiner and Norman 2008). In the case of functional status or symptoms there is no gold standard and therefore construct validity must be established in its place (Amadio 2001). Construct validity refers to the degree to which outcome measurement scores are consistent with the theoretic construct being measured (Streiner and Norman 2008). Evidence accumulates either in support or opposition of the construct validity of a specific measure. Three subtypes of construct validity (convergent, divergent and structural validity) of the PRWHE were evaluated with 122 patients post arthroplasty for osteoarthritis of the first carpometacarpal joint (MacDermid et al 2007). Convergent validity refers to how well the scores from a measure correlate with scores from other similar measures (Streiner and Norman 2008). The study by MacDermid et al (2007) supported convergent validity of the PRWHE in this post arthroplasty population when compared with the Australian/Canadian Osteoarthritis Hand Index (AUSCAN) and the Disability of the Arm, Shoulder and Hand (DASH). Divergent validity, refers to the degree with which the scores from a scale do not correlate to scores from a scale that measures a dissimilar construct (Streiner and Norman 2008). This was demonstrated by a lack of association, as predicted, with self-reported hand appearance. Structural validity refers to the degree to which the scores of a scale are an adequate reflection of the dimensionality of the construct (Streiner and Norman 2008). The findings of MacDermid et al (2007) challenged the structural validity of the PRWHE as both the pain and function subscales unexpectedly loaded onto one factor. If subscales are valid and distinct

they should load on different factors (Thompson 2002). The MacDermid et al (2007) study had well defined hypotheses and utilised appropriate statistical analyses to test them, including specification of the direction and magnitude of correlations to be examined a priori. However, it received a fair rating for methodological quality due to not adequately reporting how missing data were dealt with. Generalisability of the results was good with a thorough description of the sample characteristics.

While the study by MacDermid et al (2007) utilised Classical Test Theory (CTT) to assess validity of the PRWHE, Rasch analysis, a form of Item Response Theory (IRT), may also be used in the development and evaluation of the internal construct validity of a patient reported outcome measure (Tennant et al 2004, Kersten and Kayes 2011). Internal construct validity is confirmed when a scale conforms to the definition of the construct. Packham and MacDermid (2013) utilised Rasch analysis to assess the psychometric properties of the PRWHE in 264 patients with acute traumatic or postoperative wrist and/or hand injuries. These authors found that while the PRWHE does not fit the Rasch model when considered as a whole, three subscales demonstrate a good fit to the Rasch model if the ability scale is separated into two subscales of specific and usual activities. Tests for unidimensionality of the entire scale also failed. However, acceptable unidimensionality was demonstrated for the three subscales, indicating that each subscale represents a unique construct (Packham and MacDermid 2013).

The Rasch analysis also identified disordered thresholds; i.e. increases in the trait (e.g. pain) did not correspond with ordering of the response categories. Therefore, it was necessary to collapse scoring categories for six of the 15 items. For example, response categories for the work item were re-scored to create six ordered categories. Results of the Rasch analysis suggest that for comparisons to be made across populations six of the 15 items need to be rescored which without the aid of a specialised computer programme would be laborious and error prone in a clinical setting (Packham and MacDermid 2013). The study by Packham and MacDermid (2013) had an appropriate sample size and met all but one of the COSMIN criteria for internal validity: it did not adequately describe how missing data were dealt with, along with the method of estimation used. A thorough description of sample demographics was given and a wide range of wrist and hand conditions were included in the sample ensuring good generalisability of the results.

A process of cross-cultural adaptation, involving culturally appropriate adaptation of questions if needed and language translation, must be followed if a patient reported outcome measure is to be administered to a different culture (Beaton et al 2000). The PRWHE has been cross-culturally adapted into Dutch and Italian (Brink et al 2009, Fairplay et al 2012). Both of these studies concluded the adapted instruments were valid and reliable. However, they achieved overall ratings of poor for methodological quality. A key limitation of both studies was a very small sample size, especially for a process of cross-cultural adaptation. Additionally, Brink et al (2009) did not present an adequate description of the translation process or the translator's expertise and the final translation was not reviewed by committee. Ease of patient comprehension was also not assessed. With ten percent of the sample failing to fully complete the questionnaire this may have been a reflection of poor comprehension. In contrast, Fairplay et al (2012) followed a

Table 3: Summary of studies examining psychometric properties of the PRWHE

Authors	Methods	Results
MacDermid and Tottenham (2004)	60 patients with wrist &/or hand problems completed the PRWHE & the DASH on their first clinic appointment & 3 months later.	Responsiveness: Demonstrated a large treatment effect (SRM: 1.51 & ES: 1.61). Slightly more responsive than the DASH (SRM: 1.51 vs 1.37).
MacDermid et al (2007)	122 patients 9-117 months post tendon interposition arthroplasty for OA of the first carpometacarpal joint completed the PRWHE, AUSCAN, DASH & SF-36.	Construct Validity: Convergent Validity: High correlations between similar subscales ($r > 0.75$). Divergent Validity: No correlation ($r < 0.05$) with self reported hand appearance. Discriminative Validity: does not discriminate between patients with localised hand OA versus those with OA affecting additional joints. Structural Validity: Pain (0.903) & function (0.906) subscales loaded on a single factor.
Brink et al (2009)	Dutch language versions of the PRWHE & DASH were completed by 58 patients with wrist &/or hand problems & then 2 days later.	Reliability: Internal consistency: Excellent for the total scale & the subscales ($\alpha = 0.89-0.95$) Test-retest reliability: High correlations for the total scale & the subscales (ICC=0.88-0.89) Construct Validity: Convergent Validity: High correlation with the DASH ($r = 0.84$)
Fairplay et al (2012)	Italian versions of the PRWHE, DASH & SF-36 were completed by 63 patients with stable wrist &/or hand problems & then 5-7 days later.	Reliability: Internal consistency: Excellent for the total scale ($\alpha = 0.96$). Test-retest reliability: Very high correlation (ICC=0.92) Construct Validity: Convergent validity: High correlation with the DASH ($r = 0.80$)
Packham and MacDermid (2013)	Rasch analysis was conducted on PRWHE scores in 264 patients with acute traumatic or postoperative wrist &/or hand injuries	Rasch Model: PRWHE does not fit the Rasch Model as a whole (item-trait interaction $\chi^2(30) = 83.15$, $p < 0.001$) Three subscales of pain, specific activities & usual activities fit the Rasch Model (PSI=0.86, $\chi^2(6) = 10.01$, $p = 0.12$) Reliability: Internal consistency: Good to excellent (PSI=0.86-0.95) ($\alpha = 0.96$) Construct Validity: Unidimensionality: Entire scale failed ($p = 0.216$) Acceptable for 3 subscales: pain ($p = 0.075$ [95% CI: 0.048-0.102]), specific activities ($p = 0.076$ [95% CI: 0.048-0.103]), usual activities ($p = 0.049$ [95% CI: 0.021-0.076]) Systematic Biases: No evidence for systematic bias, however, gender, age, diagnosis & duration since injury influenced scoring whereas hand dominance & affected side did not.

Australian/Canadian Osteoarthritis Hand Index (AUSCAN), Chi square value (χ^2), Chronbach's Alpha (α), Confidence Interval (CI), Disability of the Arm, Shoulder and Hand (DASH), Effect size (ES), Intraclass correlation coefficient (ICC), Osteoarthritis (OA), Patient Rated Wrist and Hand Evaluation (PRWHE), Person Separation Index (PSI), Standardised response mean (SRM),

standardised cross-cultural adaptation process as recommended by Beaton et al (2000). The authors gave a thorough description of the translation process including; the expertise of the translators, how differences of opinion were resolved, and

the committee review process. Ease of comprehension was also assessed. However, generalisability of these results was compromised by not reporting on sample demographics such as age, gender and disease characteristics.

Table 4: Methodological quality of studies examining the psychometric properties of the PRWHE according to the COSMIN checklist

Authors	IRT model requirements	Internal consistency	Reliability	Structural validity	Hypothesis testing	Cross-cultural validity	Responsiveness	Interpretability	Generalisability	Overall rating of methodological quality
MacDermid and Tottenham (2004)							Fair	0/7	5/8	Fair
MacDermid et al (2007)				Fair	Fair			2/7	4/8	Fair
Packham and MacDermid (2013)	Good	Good		Good				4/7	7/8	Good
Brink et al (2009)		Poor	Fair		Poor	Poor		2/7	4/4	Poor
Fairplay et al (2012)		Poor	Fair		Poor	Poor		3/7	3/8	Poor

Note: Only psychometric properties evaluated to date for the PRWHE are included in the table

Internal consistency of the PRWHE

Internal Consistency, a type of reliability, refers to the homogeneity of scale items and therefore the extent to which they measure different aspects of the same construct (Giladi and Chung 2013, Streiner and Norman 2008). Good to excellent internal consistency of the PRWHE scale and its subscales was demonstrated with Person Separation Index (PSI) values ranging from 0.86 to 0.95 (Packham and MacDermid 2013). PSI statistics are similar to Cronbach's Alpha, which are used in studies employing Classical Test Theory approaches to outcome measurement. This study received a rating of good for internal consistency, although it did not adequately report on missing data. Brink et al (2009) and Fairplay (2012) found the internal consistency for the Dutch and Italian language versions of the total scale and its subscales to be excellent with Cronbach's Alpha ranging from 0.89 to 0.96. However, both of these studies received a poor rating of methodological quality with respect to internal consistency.

Reliability of the PRWHE

Test-retest reliability refers to the extent to which repeated measurement scores are the same in patients who have not changed (Streiner and Norman 2008). Brink et al (2009) and Fairplay et al (2012) observed high to very high intraclass correlation coefficients in the Dutch and Italian language versions of the PRWHE, ranging from 0.88-0.92. However, both of these studies received a fair rating of methodological quality for test-retest reliability.

Responsiveness of the PRWHE

Responsiveness refers to the ability of an outcome measure to detect change in the construct being measured over time (Streiner and Norman 2008). It is a measure of longitudinal validity (van de Ven-Stevens et al 2009). The responsiveness

of the PRWHE versus the DASH was evaluated in a cohort of 60 patients with wrist and/or hand problems (MacDermid and Tottenham 2004). The study included acute trauma and surgical patients with a variety of diagnoses such as fractures, carpal instabilities, osteoarthritis, tendon lacerations, and palmar fasciectomy. Both the PRWHE and the DASH demonstrated a large treatment effect, in both the wrist and hand injury groups following hand therapy intervention. Responsiveness was slightly higher in the PRWHE and the change measured in the two instruments was strongly correlated. This, in addition to the fact that the PRWHE contains half as many items as the DASH, led the authors to conclude that the PRWHE is a more efficient measure for detecting clinical improvement. However, hypotheses were either not formulated or they were not reported on, resulting in a rating of fair for overall methodological quality for this study. Key strengths of this study were a good sample size and a thorough description of an appropriate comparator instrument, i.e. the DASH, including its measurement properties. However, generalisability of these results was compromised due to inadequate descriptions of diagnoses, duration of time since injury/surgery, and the type of hand therapy received.

DISCUSSION

This review identified five papers designed to examine psychometric properties of the PRWHE. There was good evidence for the internal reliability of the PRWHE. However, evidence for test-retest reliability and the responsiveness of the tool was limited due to poor and fair ratings respectively for methodological quality of the relevant studies. There was also limited evidence for the convergent and divergent validity of the tool. For example, the PRWHE was compared to the AUSCAN and DASH only and not to other commonly used scales such as

the Michigan Hand Questionnaire. Evidence for the structural validity was mixed, with one study suggesting all items load onto one factor (MacDermid et al 2007) and another that it consists of three subscales (Packham and MacDermid 2013). In addition, the response categories do not appear to work as intended.

A study by MacDermid and Tottenham (2004) found that hand therapists favoured routine use of the PRWHE over the DASH due to it being quicker and easier for patients to complete and for therapists to administer and score. The cost associated with administering and analysing the PRWHE is minimal and there is no requirement for further training or specialised equipment. However, there are clinical implications based on the findings of this review for clinicians using the PRWHE. Clinical utility of the tool is challenged by the requirement for collapsing of response categories as found by Packham and MacDermid (2013). Furthermore, these authors did not provide a scoring algorithm to convert the ordinal raw PRWHE scores into interval level data. This would have been possible to do following the Rasch analysis. Until such a conversion table is available clinicians and researchers must exercise caution when considering the total summed PRWHE scores, since they remain ordinal in nature (Kersten and Kayes 2011).

Clinicians using the tool should also be aware that evidence for the content validity of the PRWHE is not fully established. For example, some patients may not be able to respond to all items since they may not have performed these activities. This is problematic both from a measurement standpoint and for people completing the tool. In addition, the items include words that may not be appropriate to a New Zealand context. Examples are the use of pounds (lb) for the weight of an object and terms such as bathroom tissue instead of toilet paper.

Limitations of this review include the quality scoring of studies by one, rather than multiple reviewers (JT). However, any uncertainties in the scoring were discussed with the second author. Strengths of this review include a thorough literature search across multiple databases and the use of an appropriate assessment tool, with established reliability and validity, to rate the methodological quality of included studies.

CONCLUSIONS

The methodological quality of the included studies was mixed, with the key limitations including a lack of reporting on how missing data were handled and in some instances small sample sizes. Four studies that utilised CTT were found to be of fair to poor methodological quality while a recent Rasch analysis was found to be of good methodological quality. Whilst evidence for the validity and reliability of the PRWHE were established as part of this review further work is required to evaluate cultural validity in the New Zealand context. In addition, research is required to further evaluate the internal construct validity of the scale and make recommendations regarding possible changes needed to the response categories and subscales.

KEY POINTS

- Four studies which utilised Classical Test Theory were found to be of fair to poor methodological quality while a Rasch analysis was found to be of good methodological quality.

- Rasch analysis supports the internal validity and reliability of a three subscale PRWHE instrument (Pain, Specific Activities, and Usual Activities).
- Clinicians should exercise caution when considering total summed scores.
- Cultural validity of the PRWHE in the New Zealand context needs to be evaluated.

SOURCES OF FUNDING

No financial support was received for this study.

ADDRESS FOR CORRESPONDENCE

Julie Taylor, Whangarei Physiotherapy Services, 445 Kamo Road, Whangarei, 0112, Ph: 09 459 5000, 027 841 6030, julie.taylor@windowslive.com

REFERENCES

- Accident Compensation Corporation (2009) Guide to outcome measure reporting, Wellington, New Zealand.
- Amadio PC (2001) Outcome assessment in hand surgery and hand therapy: An update. *Journal of Hand Therapy* 14: 63-67.
- Beaton DE, Bombardier CM, Guillemin FM, Ferraz MB (2000) Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 25: 3186-3191.
- Brink S, Voskamp EG, Houpt P, Emmelot CH (2009) Psychometric properties of the patient rated wrist/hand evaluation - Dutch language version (PRWH/E-DLV). *Journal of Hand Surgery: European Volume* 34: 556-557.
- Fairplay T, Atzei A, Corradi M, Luchetti R, Cozzolino R, Schoenhuber R (2012) Cross-cultural adaptation and validation of the Italian version of the patient-rated wrist/hand evaluation questionnaire. *Journal of Hand Surgery: European Volume* 37: 863-870.
- Giladi A M, Chung KC (2013) Measuring outcomes in hand Surgery. *Clinics in Plastic Surgery* 40: 313-322.
- Hoang-Kim A, Pegreff F, Moroni A, Ladd A (2011) Measuring wrist and hand function: Common scales and checklists. *Injury* 42: 253-258.
- Holmbeck GN, Devine K A (2009) Editorial: an author's checklist for measure development and validation manuscripts. *Journal of Pediatric Psychology* 34: 691-696.
- Horner D, Larmer PJ (2006) Health outcome measures. *New Zealand Journal of Physiotherapy* 34: 17-24.
- Jette DU, Halbert J, Iverson C, Miceli E, Shah P (2009) Use of standardized outcome measures in physical therapist practice: perceptions and applications. *Physical Therapy* 89: 125-135.
- Kersten P, Kayes NM (2011) Outcome measurement and the use of Rasch analysis, a statistics-free introduction. *New Zealand Journal of Physiotherapy* 39: 92-99
- MacDermid JC (1996) Development of a scale for patient rating of wrist pain and disability. *Journal of Hand Therapy* 9: 178-183.
- MacDermid JC, Tottenham V (2004) Responsiveness of the disability of the arm, shoulder, and hand (DASH) and patient-rated wrist/hand evaluation (PRWHE) in evaluating change after hand therapy. *Journal of Hand Therapy* 17: 18-23.
- MacDermid JC, Wessel J, Humphrey R, Ross D, Roth JH (2007) Validity of self-report measures of pain and disability for persons who have undergone arthroplasty for osteoarthritis of the carpometacarpal joint of the hand. *Osteoarthritis And Cartilage* 15: 524-530.
- Mijnarends DM, Meijers JM, Halfens RJ, ter Borg S, Luiking YC, Verlaan S, Schoberer D, Jentoft AJ, van Loon LJ, Schols JM (2013) Validity and reliability of tools to measure muscle mass, strength, and physical performance in community-dwelling older people: A systematic review. *Journal of the American Medical Directors Association* 14: 170-178.
- Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC (2010a) The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health

- status measurement instruments: an international Delphi study. *Quality of Life Research* 19: 539-549.
- Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC (2010b) The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology* 63: 737-745.
- Mokkink LB, Terwee CB, Gibbons E, Stratford PW, Alonso J, Patrick DL, Knol DL, Bouter LM, de Vet HC (2010c) Inter-rater agreement and reliability of the COSMIN (Consensus-based standards for the selection of health status measurement instruments) checklist. *BMC Medical Research Methodology* 10.
- Nunnally JC, Bernstein IH (1994) *Psychometric theory* (3rd edn). New York: McGraw-Hill.
- Packham T, MacDermid JC (2013) Measurement properties of the patient-rated wrist and hand evaluation: Rasch analysis of responses from a traumatic hand injury population. *Journal of Hand Therapy* 26: 216-224.
- Rothman ML, Beltran P, Cappelleri JC, Lipscomb J, Teschendorf B (2007) Patient-reported outcomes: conceptual issues. *Value in Health* 10: S66-S75.
- Streiner DL, Norman GR (2008) *Health Measurement Scales: A practical guide to their development and use* (4th edn). Oxford: Oxford University Press.
- Tennant A, McKenna SP, Hagell P (2004) Application of Rasch analysis in the development and application of quality of life instruments. *Value in Health* 7: S22-S26.
- Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC (2012) Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research* 21: 651-657.
- The Physiotherapy Board of New Zealand (2009) *Physiotherapy competencies for physiotherapy practice in New Zealand*, Wellington, New Zealand.
- Thompson B (2002) *Exploratory and confirmatory factor analysis. Understanding concepts and applications*. Washington: American Psychological Association.
- van de Ven-Stevens LA, Munneke M, Terwee CB, Spauwen PH, van der Linde H (2009) Clinimetric properties of instruments to assess activities in patients with hand injury: a systematic review of the literature. *Archives of Physical Medicine and Rehabilitation* 90: 151-169.