International Orthopaedics (SICOT) (2008) 32:497–504
DOI 10.1007/s00264-007-0344-7

ORIGINAL PAPER

Critical analysis of outcome measures used in the assessment of carpal tunnel syndrome

Senthil Nathan Sambandam · P. Priyanka · Arif Gul · Balakrishnan Ilango

Received: 23 October 2006 / Revised: 16 January 2007 / Accepted: 24 January 2007 / Published online: 17 March 2007

Abstract Clinicians and researchers are confounded by the various outcome measures used for the assessment of carpal tunnel syndrome (CTS). In this study, we critically analysed the conceptual framework, validity, reliability, responsiveness and appropriateness of some of the commonly used CTS outcome measures. Initially, we conducted an extensive literature search to identify all of the outcome measures used in the assessment of CTS patients, which revealed six different carpal tunnel outcome measures [Boston Carpal Tunnel Questionnaire (BCTQ), Michigan Hand Outcome Questionnaire (MHQ), Disability of Arm, Shoulder and Hand (DASH), Patient Evaluation Measure (PEM), clinical rating scale (Historical-Objective (Hi-Ob) scale) and Upper Extremity Functional Scale (UEFS)]. We analysed the construction framework, development process, validation process, reliability, internal consistency (IC), responsiveness and limitations of each of these outcome measures. Our analysis reveals that BCTQ, MHQ and PEM have comprehensive frameworks, good validity, reliability and responsiveness both in the hands of the developers, as well as independent researchers. The UEFS and Hi-Ob scale need validation and reliability testing by independent researchers. Region-specific measures like DASH have good frameworks and, hence, a potential role in the assessment of CTS but they require more validation in exclusive carpal tunnel patients.

Résumé Les cliniciens et les chercheurs sont submergés par le nombre de mesures utilisées pour l'évaluation du syndrome du canal carpien (CTS). Nous avons réalisé pour cette étude une étude critique de ces différentes mesures au travers de la littérature. Six différentes mesures sont utilisées : le questionnaire de Boston (BCTQ), le questionnaire de Michigan (MHQ), le score DASH, l'évaluation PEM, le questionnaire Hi-Ob scale et la mesure UEFS. Nous avons étudié sur tous les plans ces différents scores. Les mesures de type BCTQ, MHQ et PEM sont fidèles et utiles aussi bien pour les chirurgiens que pour les chercheurs. Les mesures UEFS et Hi-Ob scale nécessitent une validation par des examinateurs indépendants, la technique DASH nécessite une validation pour les patients présentant une lésion isolée du canal carpien.

S. N. Sambandam
Wrightington Hospital,
Wigan, UK

P. Priyanka
Royal Albert Edward Infirmary,
Wigan, UK

A. Gul
Swansea, UK

B. Ilango
Fairfield General Hospital,
Bury, UK

S. N. Sambandam (✉)
24 Dobson Close,
Wigan WN6 9ES, UK
e-mail: sam_senthil2002@yahoo.co.in

Introduction

Carpal tunnel syndrome (CTS) is a common hand disorder, with prevalence varying from 5.8% in females to 0.6% in males. Aetiology, pathogenesis, clinical presentation, diagnosis and treatment options have been extensively studied. Various outcome measures are used to assess the disability associated with CTS, to guide treatment decision, to assess the success of treatment postoperatively and also to compare the preoperative disability and treatment effects

within the same patient, as well as between two groups. Traditionally, clinical symptoms and signs were used as outcome measures. Some studies used generic quality of life measures like SF-36 to assess the outcome after carpal tunnel release [1, 7]. Levine et al. [18] in 1993 introduced the first disease-specific questionnaire, the Boston Carpal Tunnel Questionnaire (BCTQ), to assess the outcome after carpal tunnel release. Since then, various other upper-limb outcome measures have been used in the assessment of carpal tunnel patients. Some of the commonly used outcomes measures, identified by our literature search, are Disability of Arm, Shoulder and Hand (DASH) [8, 10, 11, 15, 22], clinical rating system (Hi-Ob scale) [9, 20], Patient Evaluation Measure (PEM) [6, 12, 19], Michigan Hand Outcome Questionnaire (MHQ) [3, 4, 16] and Upper Extremity Functional Scale (UEFS) [21]. In this article, we studied the development process, the validation process, reliability, internal consistency (IC) and the sensitivity testing process of each outcome measure used in CTS and we also evaluated the advantages, disadvantages and appropriateness of some of the outcome measures used for the assessment of patients with CTS.

Materials and methods

In this study, we conducted an extensive literature search to identify all of the outcome measures used in the assessment of patients with CTS. We restricted our search to reports in the English literature. We searched Medline, EMBASE and CINAHL using the OVID interface. We also searched other interfaces like EBSCO, Academic Search Premier and Proquest, as well as individual publishers' websites such as MDconsult, Science Direct, Lippincott Wilkins and Williams (LWW), Lippincott Ravens, Blackwell Synergy, Wiley Interscience, Taylor and Francis and Springerlink. We used search terms such as "carpal tunnel," "outcome measures," "outcome scale," "assessment scale," "disease-specific questionnaire," "patient-generated questionnaire" and "patient-oriented questionnaire" either alone or in different combinations. Our search identified six outcome measures that are commonly used in the assessment of patients with CTS. We then critically evaluated the construction framework, development and validation process, reliability and sensitivity assessment, appropriateness and limitations of each outcome measure.

Factors to be considered when selecting an outcome measure for any disorder

Different authors like Bindra et al. [2], Szabo [24] and Schuind et al. [23] have reviewed various aspects of outcome measures used in hand surgery. Schuind et al. [23] reviewed the various outcome measures used for hand and upper-extremity disorders and emphasised the need for a comprehensive outcome assessment process including objective, subjective and laboratory criterion. Bindra et al. [2], in their review, highlighted the importance of considering factors like the stages in the development of a questionnaire, attributes of a questionnaire and factors that affect filling the questionnaire. Based on these studies, patients' input and our own experience, we listed the factors which are required for the selection of an outcome measure (Table 1). We analysed the outcome measures used in CTS based on these variables and have given some suggestions based on our analysis.

Item generation and item reduction

An ideal outcome measure should have the item generation process clearly described, involve experts, patients and previous literature in the item generation process, clearly explain the item reduction process and should have performed a factor analysis to verify its factor structure.

Variables measured

An ideal outcome measure should include patient-oriented outcome variables, such as pain, symptoms scale, function scale, satisfaction scale and objective variables including clinical signs, physical examination findings, experimental and laboratory tests.

Type of outcome measure

Ideally, the outcome should be reported with one generic measure to compare with the disability and quality of life associated with other disorders, one region-specific measure

Table 1 Variables to be considered when selecting an outcome measure

Variables
1. Generic, disease-specific or region-specific
2. Patient-oriented, clinical, lab-based or combined
3. Item generation—patient-generated, physician-generated or both
4. Patient-filled, physician-filled or combined
5. Validity established for that particular disease (content, criterion, construct, cross-cultural, factor analysis and longitudinal validity)
6. Reliable (IC and test–retest reliability)
7. Sensitiveness or responsiveness
8. Easy to fill with less assistance
9. Less time consuming
10. Acceptable
11. Easily available with no copyright or cost restrictions
12. Quantitative and amenable to statistical analysis
13. Standardised for the disease and normal population

to enable comparison with disease of the same region and one disease-specific measure to enable comparisons of outcome following different treatments and to identify subtle alterations in the disease-related health status and quality of life.

Validation

An ideal outcome measure must have acceptable face validity, content validity, construction validity, criterion validity and longitudinal validity. They should show similar validation results in the hands of the developers and also in the hands of independent researchers (generalisability). The outcome should have good cross-cultural validity.

Reliability

An ideal outcome measure should have very good test–retest reliability and a well established memory recall period. The reliability is considered good if it is statistically significant or, according to Landis and Koch [17], a Kappa of 0.21 to 0.40 is considered to be fair, 0.41 to 0.60 moderate, 0.61 to 0.80 substantial and 0.81 to 1.0 as excellent. Similarly, the items should be internally consistent with a minimum Cronbach's alpha of 0.6.

Responsiveness

An outcome measure should have a high standardised response mean (SRM) and high effect size. According to Cohen, an SRM of 0.2 is small, 0.5 is medium and 0.8 is large [5].

Other aspects of an ideal outcome measure

An outcome measure should be free of cost and easy to fill in without missing too many questions with good response rate, and the time needed to fill the questionnaire in should be clearly established and it should be appropriate in clinical (outpatient, inpatient) and research settings. The outcome variable should have a final quantitative value which can be used for statistical comparisons. Above all, the outcome measure should have high acceptability among clinicians and researchers.

Boston Carpal Tunnel Syndrome Questionnaire (BCTQ)

Conceptual framework

BCTQ is the most commonly used outcome measure in the assessment of patients with CTS. Levine et al. [18] developed this questionnaire in 1993. This is a disease-specific patient-filled questionnaire. Both physicians and patients were involved in the item generation process. Domains in the questionnaire were decided by the patients and physicians, and factorial analysis was not performed by the developers to check the domain structure of the questionnaire. The questionnaire comprises two scales, a symptom severity scale and a functional status scale (Table 2). The symptom severity scale has 11 questions scored from 1 point (mildest) to 5 points (most severe). The similarly functional status scale has eight questions scored from 1 point (no difficulty with activity) to 5 points (cannot perform the activity at all). The overall score for both scales was calculated as the mean of the items. The scale was tested on 38 patients who had surgical treatment of CTS.

Validity

The developers checked the criterion validity by comparing the score with grip strength, pinch strength, two-point discrimination and pressure sensitivity, and this correlated well with these variables in the expected direction. Cross-cultural validation has been performed in the Italian, Swedish and Portuguese languages.

Reliability

The scales were highly reproducible and internally consistent in the hands of the developer, with a reproducibility of $r=0.91$ for symptom severity scale and $r=0.93$ for functional status scale IC (Cronbach's alpha) of 0.89 for the symptom severity scale and 0.91 for the functional status scale.

Responsiveness

The developers noticed very good responsiveness with an effect size (ES) of 1.4 for the symptom severity scale and 0.82 for the functional scale. Similarly, high responsiveness was shown by other researchers, such as Greenslade et al. [10] (SRM of 1.02 for the symptom scale and 0.62 for the functional scale at 3 months) and Gay et al. [7] (SRM/ES of 1.21/1.30 and 1.66/1.71, 0.46/0.48 and 1.05/1.05, 1.67/1.74 and 2.01/1.96 for the whole questionnaire, symptom scale and functional scale at 6 weeks and 12 weeks, respectively). They also demonstrated that the questionnaire does not have ceiling or floor effects.

Limitations

Even though the reliability, responsiveness and criterion validity has been well established over the years, the construction and domain structure were never validated, either by the developers or by independent researchers. Similarly, this questionnaire was not validated in patients claiming for workers compensation, thereby, precluding its

Table 2 Boston Carpal Tunnel Syndrome Questionnaire (BCTQ) [18], clinical rating scale (Hi-Ob scale) [9], Patient Evaluation Measure (PEM) [19]

Boston Carpal Tunnel Syndrome Questionnaire
Boston Carpal Tunnel Syndrome Questionnaire (BCTQ)
Symptom severity scale (11 items)
1. How severe is the hand or wrist pain that you have at night?
2. How often did hand or wrist pain wake you up during a typical night in the past two weeks?
3. Do you typically have pain in your hand or wrist during the daytime?
4. How often do you have hand or wrist pain during daytime?
5. How long on average does an episode of pain last during the daytime?
6. Do you have numbness (loss of sensation) in your hand?
7. Do you have weakness in your hand or wrist?
8. Do you have tingling sensations in your hand?
9. How severe is numbness (loss of sensation) or tingling at night?
10. How often did hand numbness or tingling wake you up during a typical night during the past two weeks?
11. Do you have difficulty with the grasping and use of small objects such as keys or pens?
Functional status scale (8 items)
1. Writing
2. Buttoning of clothes
3. Holding a book while reading
4. Gripping of a telephone handle
5. Opening of jars
6. Household chores
7. Carrying of grocery basket
8. Bathing and dressing
Clinical rating scale (Hi-Ob scale)
(A) any kind of paraesthesia in the hand (numbness, tingling, burning, etc.) with regard to its temporal onset and duration
(B) sensory function in the median nerve distribution of the hand
(C) motor function of median innervated muscles of the hand
(D) trophism of the thenar eminence
(E) pain, reported as dull or aching discomfort, in the hand, forearm or upper arm
Patient Evaluation Measure (PEM)
Part One: Treatment
Please put a circle around the number that is closest to the way you feel about how things have been for you. There are no right or wrong answers.
1. Throughout my treatment I have seen the same doctor:
2. When the doctor saw me, he or she knew about my case:
3. When I was with the doctor, he or she gave me the chance to talk:
4. When I did talk to the doctor, he or she listened and understood me:
5. I was given information about my treatment and progress:
Part Two: How Your Hand is Now?
1. The FEELING in my hand is now:
2. When my hand is cold and/or damp, the PAIN is now:
3. Most of the time, the PAIN in my hand is now:
4. When I try to USE my hand for fiddly things, it is now:
5. Generally, when I MOVE my hand it is:
6. The GRIP in my hand is now:
7. For everyday ACTIVITIES, my hand is now:
8. For WORK, my hand is now:
9. When I look at the appearance of my hand now, I feel:
10. Generally, when I think about my hand, I feel:
Part Three: Overall Assessment
1. Generally, my treatment at the hospital has been:
2. Generally, my hand is now:
3. Bearing in mind my original injury or condition, my hand is now:
Are there any other comments you wish to make?
Thank you very much indeed for your help

use in this set of patients. The time required to fill in this questionnaire is not clearly established and nor are the commonly missed items not mentioned by developers or independent researchers. Despite these limitations, this questionnaire continues to be the most commonly used in the assessment of patients with CTS.

Upper Extremity Functional Scale (UEFS)

Conceptual framework

This scale was developed in 1997 by Pransky et al. [21]. It is a region-specific questionnaire developed to assess the outcome of work-related upper extremity disorders. This scale is a patient-filled questionnaire measuring the functional component of physical health and has eight items derived from a pool of 12 items generated by physicians, patients and occupational therapists. The item responses were on a 1–10 response scale for CTS, with 1 being no problem with that particular activity and 10 being cannot do it at all (Table 3). This scale was tested on 165 patients with CTS.

Validity

Principal component factor analysis revealed that 42% to 56% of the inter-item variance was explained by a single factor. Also, no other factor was found with high Eigen values. Validity was also established by demonstrating varying UEFS scores for patients with varying duration of symptoms and the impact of illness on working status. It was validated further by demonstrating a high degree of correlation between the UEFS score and other continuous measures of self-reported pain, psychological measure and physical examinations.

Reliability

The IC was found be excellent for patients with CTS, with alpha ranging from 0.83 to 0.93 for the various items.

Table 3 Upper Extremity Functional Scale (UEFS) [21]

	(No problem)	(Major problem)
Sleeping	1	10
Writing	1	10
Opening jars	1	10
Picking up small objects with fingers	1	10
Driving a car more than 30 minutes	1	10
Opening a door	1	10
Carrying milk jug from the refrigerator	1	10
Washing dishes	1	10

Responsiveness

High responsiveness was established by showing a high correlation between change in UEFS to change in other measures of disease and also by demonstrating variable SRM in surgical subgroups as compared to non-surgical patients.

Other features

The authors also demonstrated the absence of floor effects. The scale is readily available electronically and its use for scientific and clinical purposes is unrestricted.

Limitations

The scale measures only the functional component of physical health and does not measure the symptoms nor objective components of physical, social and mental health. This scale has not gained wide acceptance for the assessment of CTS. It has also not undergone cross-cultural validation. Its validity, reliability and responsiveness has not been checked by independent researchers, apart from the developers. The time to fill in this questionnaire, the degree of assistance needed from the physician and the questions commonly missed are not mentioned. It is not standardised for a normal population and the influences of age, worker compensation issues and psychological status on the score has not be studied. Further, the items are not weighted and the test–retest reliability was never assessed.

Michigan Hand Outcomes Questionnaire (MHQ)

Conceptual framework

This questionnaire was developed by Chung et al. [4] in 1998 for the assessment of outcome for various hand disorders. The 37 items in the questionnaire were from a pool of 100 questions which was reduced by factor analysis. The 37 items are grouped into six domains: (1) overall hand function, (2) activity of daily living, (3) pain, (4) work performance, (5) aesthetics and (6) patient satisfaction with hand function. This scale was tested prospectively on a population of 200 patients with varying hand disorders, including those with CTS.

Validity

The developers also verified the construction of 37 items in six hypothetical subscales by factor analysis and

demonstrated good validity by showing good correlation between similar subscales in MHQ and SF-12 and with ADL, work performance and pain scales. Convergent validity was demonstrated by an a priori prediction model and divergent validity was demonstrated by different subscale scores for different afflictions of the hand.

Reliability

The developers showed the test–retest reliability from 0.81 to 0.97 and IC from 0.86 to 0.97 for various subscales.

Responsiveness

Kotsis and Chung [16] noted that the SRM varied from 0.5 to 0.6 for the activity subscale and from 0.9 to 1.1 for the pain and satisfaction scale, thereby, establishing its responsiveness.

Other features

This scale is electronically available and its use is unrestricted for clinical and research purpose. Both weighted and unweighted questionnaires were tested and no definite differences were found by the developer. Hence, an unweighted score was recommended for simplicity. The developer states the mean time to fill in the questionnaire was 10 minutes (range, 7–20 minutes). Comparison between outcomes for different hand disorders and different procedures for the same hand disorders is possible with this questionnaire. The MHQ was also found to be useful when independent scores from multiple domains were required. Another unique feature of this questionnaire is that four domains measure both right and left hand outcomes and, hence, comparison with an unaffected control hand is possible. The MHQ can also measure function and symptom outcomes separately. The authors admit that the generalisability of the findings of the study is not possible because of the high proportion of the female population in the study population of 50 patients.

Limitations

The questionnaire was tested in a small sample size of 200 patients with mixed hand disorders and has not been exclusively validated in patients with CTS. The scale did not measure social and mental health. Responsiveness of the MHQ was not established by the developer. The acceptability and generalisability of the MHQ is not evaluated. Floor effects and ceiling effects were not recorded. Cross-cultural validation has not been carried out. There is only one report on the use of this questionnaire for patients with CTS from independent centres.

Historical-Objective scale (Hi-Ob scale)

Conceptual framework

This is a physician-oriented scale based on clinical history and physical examination findings. Giannini et al. [9] developed this score in 2002 and validated it in 168 consecutive CTS hands. The scale was based on the parameters outlined in Table 2. Based on the presence of findings A to D, the patients were grouped into one of the five stages. The parameter E was marked as a dichotomous variable of yes or no. Hence, the final score is the combination of stage and the dichotomous variable (e.g. 2P or 3 or 4P). Criteria to group the patients into different stages based on the four parameters have been established.

Validity

The authors established validity by showing a high correlation between the Hi-Ob score and symptom and functional subscales of BCTQ and neuro-physiological measures.

Reliability and responsiveness

Reproducibility was established by demonstrating a percentage agreement of 78% and a Cohen coefficient of 0.69, and good responsiveness was also noted.

Other features

It is claimed to be both a physician- and patient-oriented programme, and better than just patient-oriented programs.

Limitations

This scale has numerous disadvantages, such as that it has never been used or validated by independent researchers, cross-cultural validation has not been done and variables like time to fill in questionnaire and commonly missed items have not been established. The scale does not cover different components of physical health (symptoms, function, ADL, dexterity, satisfaction) or social and mental health.

Patient Evaluation Measure (PEM)

Conceptual framework

Reported by Macey et al. [19] in 1995, PEM is a self administered measure of physical health (Table 2). It uses a visual analogue score and has three components: patients' opinion on the delivery of care, hand health profile and overall assessment. Nothing has been mentioned about item

generation, item reduction or other aspects of the development process.

Validity

The authors, in their initial report, described the components of the questionnaire and failed to mention anything about the development process and the validation process. Interestingly, the first report on the validation of this measure was published in the year 2000, nearly five years after the initial report, from an independent centre by demonstrating a high correlation between the grip strength and level of distress items in the questionnaire. In 2005, Hobby et al. [12] assessed its usefulness in patients with CTS and validated it by demonstrating its high correlation with the DASH score and objective measures like grip strength, sensibility and dexterity.

Reliability

The overall reliability of PEM was found to be 0.83 and the inter-item consistency was 0.88. Hobby et al. [12] also revealed a high IC of PEM (0.94).

Responsiveness

Hobby et al. [12] also reported high responsiveness, with an SRM of 0.94.

Other features

Simplicity was assessed by quantifying the help needed to fill in the questionnaire and the developers claim that the questionnaire is very simple to fill. This questionnaire is electronically available and is free.

Limitations

This measure has not been used by anyone except its developers, the study population was a non-random sample, the sample size was too low (35 patients) with no mention of patient characteristics, no cross-cultural validation was carried out and it measures only the physical components of health.

Disability of Arm, Shoulder and Hand questionnaire (DASH)

Conceptual framework

This is a region-specific scale which is a joint initiative of the American Academy of Orthopaedic Surgeons (AAOS), the Council of Musculoskeletal Specialty Societies (COMSS), and the Institute for Work and Health (Toronto, Ontario, Canada). This scale item generation and item reduction process is well described by Hudak et al. [13] in 1996. In Stage 1, Item Generation, a group of methodologists and clinical experts reviewed 13 outcome measurement scales and generated a list of 821 items. In Stage 2a, Initial Item Reduction, these 821 items were reduced to 78 items using various strategies, including the removal of items which were generic, repetitive, not reflective of disability, not relevant to the upper extremity or not relevant to one of the targeted concepts of symptoms and functional statuses. Items not highly endorsed in a survey of content experts were also eliminated. Stage 2b, Further Item Reduction, was based on field testing, in which patients completed the 78-item questionnaire and the final scale with 30 items was developed. This scale is self-administered and assesses physical and social components of health.

Validity

Hobby et al. [12] established criteria validity by demonstrating a strong correlation between DASH and PEM. Apart from this report, no other study has assessed the validity of this questionnaire in patients with CTS.

Reliability

The test–retest reliability of DASH was demonstrated by Greenslade et al. [10] (intra-class correlation, ICC=0.9).

Responsiveness

Greenslade et al. [10] recorded a responsiveness of SRM=0.66 for DASH. Gummesson et al. [11] in 2003 studied the longitudinal construct validity of the DASH questionnaire and found that the SRM was 1.0 and the effect size was 0.7 for CTS. Gay et al. [7] compared the DASH responsiveness to the responsiveness of BCTQ and SF-36 and found DASH (SRM=1.13, ES=1.01) to have a better responsiveness than SF-36 (SRM=0.52, ES=0.57). Hobby et al. [12] in 2005 evaluated the responsiveness of DASH and found an effect size of 0.49. Similarly, Kotsis [16] demonstrated moderate responsiveness, with an SRM of 0.7.

Other features

Greenslade et al. [10] established the mean time to complete the questionnaire as 6.8 min and the minimal important change for the DASH score to be 10 points. It has been validated in the French, Spanish and Japanese languages [7, 8, 14, 22].

Limitations

This questionnaire is extensively studied for other upper limb disorders but not for CTS. The IC of the questionnaire in patients with CTS, floor and ceiling effects and the commonly missed questions by carpal tunnel patients has not been established.

Discussion

It is evident from the critical analysis of some of the outcome measures used in patients with CTS that none of them can be considered to be the best or ideal outcome measure. There is scope for improvement for the outcome measures in CTS. Each measure has its own advantages and limitations. Hence, the clinician should select the outcome measure based on the primary purpose of their study. The clinician should consider other factors that will influence the choice, including the population selected, setting (hospital outpatient clinic, inpatient), mode of administration (postal or personal) and time constraints (shorter versus longer).

References

1. Atroshi I, Johnsson R, Ornstein E (1998) Patient satisfaction and return to work after endoscopic carpal tunnel surgery. J Hand Surg [Am] 23(1):58–65
2. Bindra RR, Dias JJ, Heras-Palau C, Amadio PC, Chung KC, Burke FD (2003) Assessing outcome after hand surgery: the current state. J Hand Surg [Br] 28(4):289–294
3. Chung KC, Hamill JB, Walters MR, Hayward RA (1999) The Michigan Hand Outcomes Questionnaire (MHQ): assessment of responsiveness to clinical change. Ann Plast Surg 42(6):619–622
4. Chung KC, Pillsbury MS, Walters MR, Hayward RA (1998) Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. J Hand Surg [Am] 23(4):575–587
5. Cohen J (1988) Statistical power analysis for the behavioural sciences, 2nd edn. Lawrence Earlbaum, Hillsdale, New Jersey
6. Dias JJ, Bhowal B, Wildin CJ, Thompson JR (2001) Assessing the outcome of disorders of the hand. Is the patient evaluation measure reliable, valid, responsive and without bias? J Bone Joint Surg Br 83(2):235–240
7. Gay RE, Amadio PC, Johnson JC (2003) Comparative responsiveness of the disabilities of the arm, shoulder, and hand, the carpal tunnel questionnaire, and the SF-36 to clinical change after carpal tunnel release. J Hand Surg [Am] 28(2):250–254
8. Germann G, Harth A, Wind G, Demir E (2003) Standardisation and validation of the German version 2.0 of the Disability of Arm, Shoulder, Hand (DASH) questionnaire (in German). Unfallchirurg 106(1):13–19
9. Giannini F, Cioni R, Mondelli M, Padua R, Gregori B, D'Amico P, Padua L (2002) A new clinical scale of carpal tunnel syndrome: validation of the measurement and clinical-neurophysiological assessment. Clin Neurophysiol 113(1):71–77
10. Greenslade JR, Mehta RL, Belward P, Warwick DJ (2004) DASH and Boston Questionnaire assessment of carpal tunnel syndrome outcome: what is the responsiveness of an outcome questionnaire? J Hand Surg [Br] 29(2):159–164
11. Gummesson C, Atroshi I, Ekdahl C (2003) The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: longitudinal construct validity and measuring self-rated health change after surgery. BMC Musculoskelet Disord 4:11
12. Hobby JL, Watts C, Elliot D (2005) Validity and responsiveness of the patient evaluation measure as an outcome measure for carpal tunnel syndrome. J Hand Surg [Br] 30(4):350–354
13. Hudak PL, Amadio PC, Bombardier C (1996) Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). Am J Ind Med 29(6):602–608
14. Imaeda T, Toh S, Nakao Y, Nishida J, Hirata H, Ijichi M, Kohri C, Nagano A (2005) Validation of the Japanese Society for Surgery of the Hand version of the Disability of the Arm, Shoulder, and Hand questionnaire. J Orthop Sci 10(4):353–359
15. Jester A, Harth A, Wind G, Germann G, Sauerbier M (2005) Disabilities of the arm, shoulder and hand (DASH) questionnaire: determining functional activity profiles in patients with upper extremity disorders. J Hand Surg [Br] 30(1):23–28
16. Kotsis SV, Chung KC (2005) Responsiveness of the Michigan Hand Outcomes Questionnaire and the Disabilities of the Arm, Shoulder and Hand questionnaire in carpal tunnel surgery. J Hand Surg [Am] 30(1):81–86
17. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33(1):159–174
18. Levine DW, Simmons BP, Koris MJ, Daltroy LH, Hohl GG, Fossel AH, Katz JN (1993) A self-administered questionnaire for the assessment of severity of symptoms and functional status in carpal tunnel syndrome. J Bone Joint Surg Am 75(11):1585–1592
19. Macey AC, Burke FD, Abbott K, Barton NJ, Bradbury E, Bradley A, Bradley MJ, Brady O, Burt A, Brown P (1995) Outcomes of hand surgery. British Society for Surgery of the Hand. J Hand Surg [Br] 20(6):841–855
20. Mondelli M, Ginanneschi F, Rossi S, Reale F, Padua L, Giannini F (2002) Inter-observer reproducibility and responsiveness of a clinical severity scale in surgically treated carpal tunnel syndrome. Acta Neurol Scand 106(5):263–268
21. Pransky G, Feuerstein M, Himmelstein J, Katz JN, Vickers-Lahti M (1997) Measuring functional outcomes in work-related upper extremity disorders. Development and validation of the Upper Extremity Function Scale. J Occup Environ Med 39(12):1195–1202
22. Rosales RS, Delgado EB, Diez de la Lastra-Bosch I (2002) Evaluation of the Spanish version of the DASH and carpal tunnel syndrome health-related quality-of-life instruments: cross-cultural adaptation process and reliability. J Hand Surg [Am] 27(2):334–343
23. Schuind FA, Mouraux D, Robert C, Brassinne E, Remy P, Salvia P, Meyer A, Moulart F, Burny F (2003) Functional and outcome evaluation of the hand and wrist. Hand Clin 19(3):361–369
24. Szabo RM (2001) Outcomes assessment in hand surgery: when are they meaningful? J Hand Surg [Am] 26(6):993–1002