

Patient-Reported Outcome Measures for Hand and Wrist Trauma: Is There Sufficient Evidence of Reliability, Validity, and Responsiveness?

HAND
2016, Vol. 11(1) 11–21
© American Association for
Hand Surgery 2016
DOI: 10.1177/1558944715614855
hand.sagepub.com

Peter Jonathan Dacombe¹, Rouin Amirfeyz¹, and Tim Davis²

Abstract

Background: Patient-reported outcome measures (PROMs) are important tools for assessing outcomes following injuries to the hand and wrist. Many commonly used PROMs have no evidence of reliability, validity, and responsiveness in a hand and wrist trauma population. This systematic review examines the PROMs used in the assessment of hand and wrist trauma patients, and the evidence for reliability, validity, and responsiveness of each measure in this population. **Methods:** A systematic review of Pubmed, Medline, and CINAHL searching for randomized controlled trials of patients with traumatic injuries to the hand and wrist was carried out to identify the PROMs. For each identified PROM, evidence of reliability, validity, and responsiveness was identified using a further systematic review of the Pubmed, Medline, CINAHL, and reverse citation trail audit procedure. **Results:** The PROM used most often was the Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire; the Patient-Rated Wrist Evaluation (PRWE), Gartland and Werley score, Michigan Hand Outcomes score, Mayo Wrist Score, and Short Form 36 were also commonly used. Only the DASH and PRWE have evidence of reliability, validity, and responsiveness in patients with traumatic injuries to the hand and wrist; other measures either have incomplete evidence or evidence gathered in a nontraumatic population. **Conclusions:** The DASH and PRWE both have evidence of reliability, validity, and responsiveness in a hand and wrist trauma population. Other PROMs used to assess hand and wrist trauma patients do not. This should be considered when selecting a PROM for patients with traumatic hand and wrist pathology.

Keywords: patient-reported outcomes measures, hand/wrist trauma, reliability, validity, responsiveness

Introduction

The focus of health care delivery in the 21st century is increasingly shifting, by necessity, toward achieving optimal patient outcomes at reasonable cost. This has led health care professionals, including surgeons, to consider the quality and efficacy of the interventions available to them, and how best to maximize benefit for their patients while using resources responsibly.

To accurately evaluate the efficacy of any intervention, it is vital to use appropriate outcome measures. Traditionally, hand and wrist surgeons have used objective clinical assessments, such as grip strength, range of motion, and radiographic measurements.⁹ These provide useful objective information about a given patient's outcome, but do not consider the patient's functional abilities, ability to resume normal activities of daily living or return to work, or persistent pain. "Appropriate, reliable and validated outcome measures are required that take into account all aspects of patient life that may be affected" if the outcome of hand surgery is to be fully evaluated.²²

This led to a shift in the focus of outcome analysis toward patient-reported outcome measures (PROMs), which can be a valuable tool for gathering quantitative information about the everyday experience of a patient with a given condition.⁵⁵ Longitudinal use of PROMs over multiple assessment points allows a surgeon to evaluate the progress of patient cohorts over time, as well as facilitating comparison between groups in clinical trials.³⁵ Aside from use in research, or individual surgeons critically auditing their own performance, PROMs are set to play a key role in quality measurement and resource allocation in the future. Health care purchasers are increasingly moving away from buying a procedure or service, and are moving toward

¹University Hospitals Bristol, Bristol Royal Infirmary, Bristol, England

²Queen's Medical Centre, Nottingham, England

Corresponding Author:

Peter Jonathan Dacombe, University Hospitals Bristol, Bristol Royal Infirmary, Bristol BS2 8HW, England.

Email: peter.dacombe@gmail.com

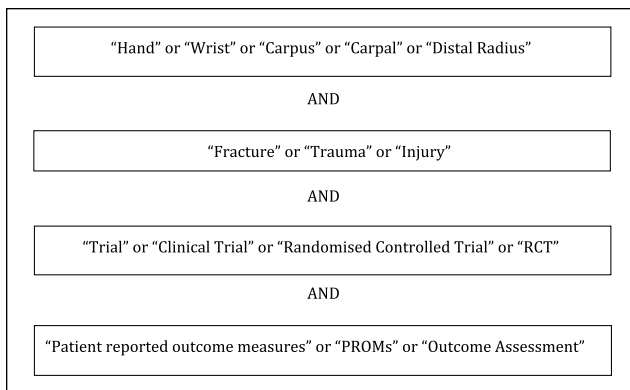


Figure 1. Keyword search terms used for primary literature search.

paying for a patient outcome. PROMs will have a key role to play in this practice.

A wide variety of outcome assessments have been proposed for the evaluation of patients with hand and wrist pathology, including generic measurements of health status, measurements specific to the upper limb, and measurements that are specific to the hand and wrist.

Generic instruments such as the Short Form 36 (SF-36)⁵³ assess the impact of musculoskeletal conditions on the overall health and well-being of a patient, and were designed for use in a broad range of conditions.⁹ They have favorable measurement properties, and can be useful for drawing comparisons about patient-reported health states across groups with varying clinical conditions.³⁵ However, the use of generic tools, such as the SF-36, in the outcome assessment of musculoskeletal conditions and interventions is of limited value because they fail to examine for condition-specific functional outcomes.

Region-specific PROMs, such as the Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire,²² claim to assess both symptoms and functional status in patient groups with upper limb musculoskeletal pathology. The DASH has been shown by multiple studies to have an excellent reliability, validity, and responsiveness profile, and it is widely used in the assessment of the upper limb. However, the DASH conceptualizes the upper extremity as a single functional unit, so it is questionable whether it is as useful an assessment tool for hand and wrist pathology as an anatomical site-specific questionnaire.

Anatomical site, or condition-specific, assessment tools include questions on aspects of health affected by the specific body region.⁵⁵ They have great potential to evaluate domains of physical function and health-related quality of life commonly affected by the specified anatomical site.³⁸ The Patient-Rated Wrist Evaluation (PRWE)²⁹ score and Michigan Hand Outcomes score¹⁰ are both examples of PROMs that are commonly used to assess outcome following hand and wrist surgery.

There are now a large number of available PROMs, even within the relatively narrow field of hand and wrist surgery. It is a natural assumption that most of the commonly used outcome assessments had their reliability, validity, and responsiveness assessed before being widely used in clinical practice. Close scrutiny of the measurement properties of these tools, however, is a relatively recent phenomenon, and many of the historical scoring systems have not undergone rigorous assessment. In addition, the characteristics of a patient undergoing elective hand and wrist surgery are likely to be very different to those who experience traumatic hand and wrist pathology. When undertaking an elective procedure for a chronic condition, a patient's function deteriorates gradually over time causing him or her to adjust one's functional requirements accordingly; in an acute traumatic event, this is not the case so a patient's ability to assess changes in function against baseline will be different. Thus, the reliability, validity, and responsiveness of an instrument cannot be assumed to be equal across each patient group.

The aims of this article are to identify which PROMs are being used in randomized controlled trials involving patients who have suffered traumatic injuries to the hand and wrist, to investigate the reported reliability, validity, and responsiveness of each PROM in a trauma population and to identify potential areas for future research.

Methods

Design

This was a systematic review of the PROMs used in randomized controlled trials to assess outcomes for hand and wrist trauma patients.

Search Strategy

Searches were performed using Medline, Pubmed, and CINAHL from the earliest records available up to the search date April 2013. The search terms used were described in Figure 1.

The lead author carried out a conventional 4-stage screening process to identify studies meeting the inclusion criteria. The senior author was available for consultation when there was doubt over the suitability of a study for inclusion. Stage 1 of the screening process removed any studies which were duplicated, stage 2 involved screening of titles, stage 3 screened abstracts, and stage 4 examined the papers in full to check they met the inclusion and exclusion criteria (Figure 2).

Study Selection

Only randomized controlled trials examining the efficacy of interventions (surgical or otherwise) in a patient group who

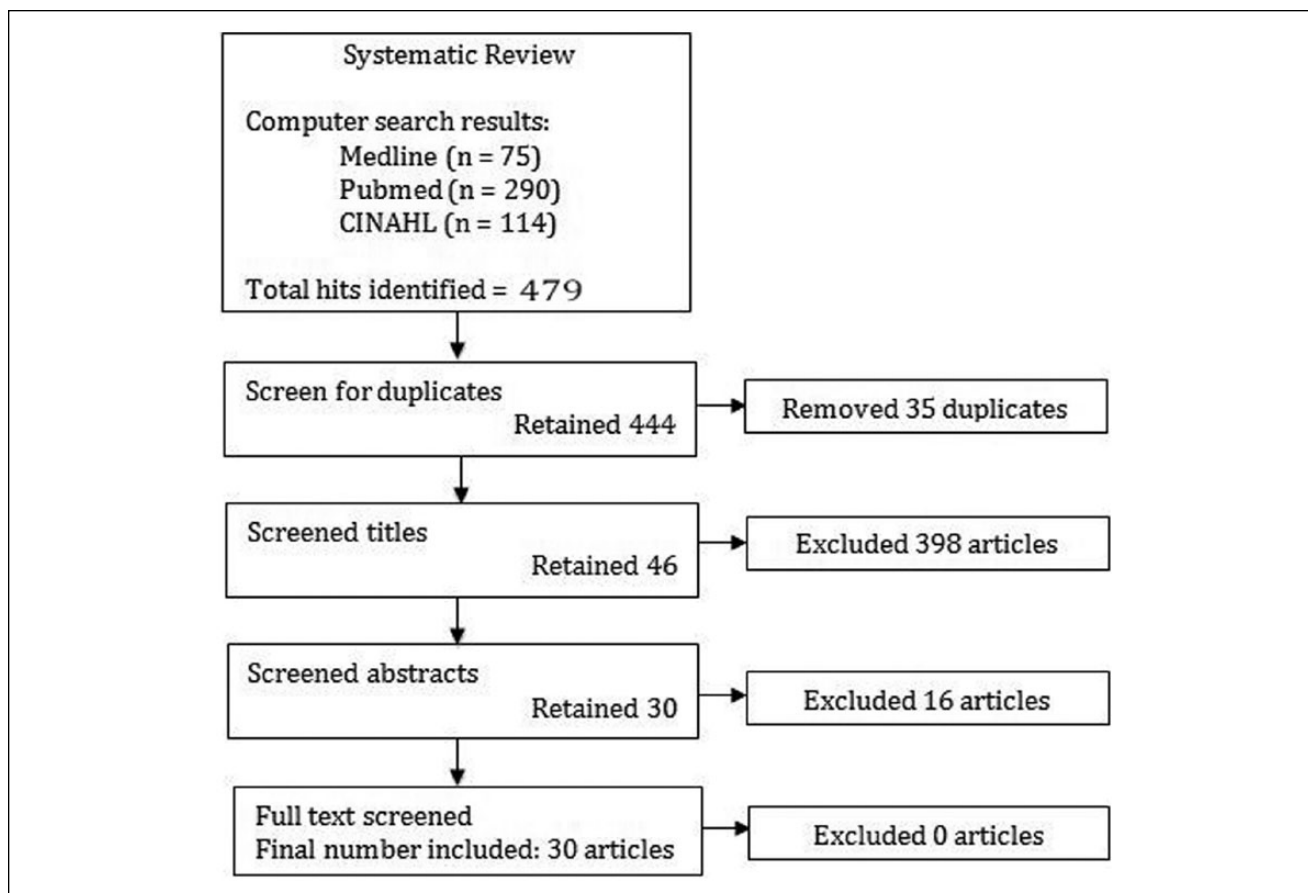


Figure 2. Summary of the 4-stage screening process and the number of papers identified/removed at each stage.

had suffered traumatic injuries to their hand or wrist were selected. Each was required to have used a PROM as part of their outcome assessment. For the purposes of this study, PROMs were defined as any outcome assessment tool that included a patient-reported component. For example the Gartland and Werley score, which is largely a surgeon-reported tool, was included on the basis of the patient-reported pain element. Studies were excluded if the patient group had suffered upper limb injuries that did not involve the hand or wrist, and if the patient group had undergone elective surgical interventions for chronic conditions. No exclusions were based on type of intervention type, length of follow-up, or numbers lost to follow-up, as these were not relevant to the article.

Data Extraction

The 30 papers accepted for inclusion in the study are documented in Table 1. For each paper, the author, year, clinical intervention, PROM, and any other outcome measure used were summarized (Table 1). The PROMs identified are shown in Table 2.

Once the relevant PROMs had been identified, a separate literature search was performed for each to investigate

whether there was evidence of reliability, validity (construct and content), and responsiveness in a hand and wrist trauma population. The search terms used for the process were “(name of patient-reported outcome measure)” AND “trauma OR injury OR fracture” AND “reliability OR validity OR responsiveness.”

All the studies identified with these searches were studied to assess whether the PROMs’ reliability, validity (construct and criterion), and responsiveness had been demonstrated in a hand and wrist trauma population. Any work examining the reliability, validity and responsiveness of the PROM in question in a hand and wrist trauma population was then thoroughly examined. If there was positive evidence of a given measurement property in a hand and wrist trauma population, this was demonstrated with +, if there was negative evidence of a given measurement property, this was demonstrated with a –, and in the event that no evidence existed on the measurement properties of a given PROM, this was represented with a 0. In the event that there was evidence of the measurement properties of a given PROM that had been gathered in a population not including hand and wrist trauma population, this was reported in brackets.

Table 1. Details of the Papers Included in the Review.

Author	Clinical population	Intervention	Patient-reported outcome measures	Other outcome assessments
Magnus et al ³¹	39 females older than 50 years with a distal radius fracture	Contralateral strength training	PRWE	Grip strength, ROM
Rostami et al ⁴¹	30 adults with hand injuries	Mirror therapy	DASH	Total active motion
Rhee et al ⁴⁰	105 adults with metacarpal fractures	Retrograde fixation using IM Kirschner wire	DASH	ROM
Schönnemann et al ⁴⁶	61 adults with distal radius fracture	External versus internal fixation	DASH, PRWE	Grip strength, radiographic assessment
Wilcke et al ⁵⁶	63 adults with distal radius fractures	Locked volar plate versus external fixation	DASH, PRWE	Radiographic assessment
Boutis et al ⁷	100 children with distal radius fractures	Cast versus splint	Activities Scale for Kids	
Belloti et al ⁶	100 adults with distal radius fractures	Percutaneous pinning versus external fixation	DASH	Pain score, ROM, grip strength, and radiographic assessment
Gruber et al ¹⁹	54 adults with distal radius fracture	Volar plate internal fixation	DASH, SF-36, Gartland and Werley	Casting score
Belloti et al ⁶	100 adults with distal radius fracture	Percutaneous pinning versus bridging external fixation	DASH	Pain score, ROM, grip strength
Pike et al ³⁹	87 adults with acute mallet finger	Volar versus dorsal versus custom thermoplastic splinting	Michigan Hand Outcomes Questionnaire	Radiographic assessment
Rozenal et al ⁴²	45 adults with distal radius fracture	Open reduction internal fixation versus closed reduction and percutaneous fixation	DASH	ROM, grip strength
Wei et al ⁵⁴	46 adults with distal radius fractures	External fixation versus radial column plate versus volar plate	DASH	Grip strength, pinch strength, ROM, radiographic assessment
Sammer et al ⁴³	99 adults with distal radius fractures	Fragment specific versus volar plate fixation	Michigan Hand Outcomes Questionnaire	ROM, radiographic assessment
Hofmeister et al ²⁰	81 patients with 5th metacarpal neck fractures	Variations of the casting technique	DASH	Grip strength, range of motion and radiographic assessment
Varitimidis et al ⁵¹	40 adults with intra-articular distal radius fractures	Fluoroscopic versus arthroscopic reduction	DASH, Mayo Wrist Score	
Lozano-Calderon et al ^{26,29}	60 patients with distal radius fracture	Early versus late mobilization	DASH, Gartland and Werley score, Mayo Wrist Score	Grip strength, ROM, radiographic assessment
Vinnars et al ⁵²	83 patients with acute scaphoid fracture	Nonoperative versus operative treatment	DASH, PRWE	
Földhazy et al ¹⁶	87 adults with distal radius fractures	No intervention	Green and O'Brien score	Grip strength, ROM, radiographic assessment
Cognet et al ¹¹	67 adults with distal radius fractures	Open reduction internal fixation	Green and O'Brien score, DASH, PRWE	Radiographic assessment

(continued)

Table 1. (continued)

Author	Clinical population	Intervention	Patient-reported outcome measures	Other outcome assessments
Schneiders et al ⁴⁵	344 patients with distal radius fractures	No intervention	Gartland and Werley score	Radiographic assessment
Atroshi et al ³	38 adults with distal radius fractures	Wrist bridging versus nonbridging external fixation	DASH	Grip strength, ROM
Müller et al ³⁶	68 patients with acute scaphoid fractures	Open reduction internal fixation	DASH, Krimmer wrist score	Pain score, radiographic assessment, grip strength
Maciel et al ³⁰	41 patients with distal radius fractures	Activity focused physiotherapy versus exercise and advice	PRWE	ROM, pain score, grip strength
Cassidy et al ⁸	323 patients with distal radius fractures	Use of Norian SRS cement	SF-36	Pain score, ROM
O'Connor et al ³⁷	68 adults with distal radius fractures	Plaster cast versus splint	Gartland and Werley score	Radiological assessment, ROM
Konrath ²³	25 patients with distal radius fractures	Trimed fragment-specific fixation	DASH, PRWE	ROM, radiographic assessment
Arora et al ²	130 adults, over 70 years, with distal radius fractures	Nonoperative treatment versus volar locking plate	DASH, PRWE, Green and O'Brien score	Pain score, radiographic assessment
Egol et al ¹⁵	280 patients with distal radius fractures	Bridging external fixation with K-wire versus volar plate	DASH	
Krukhaug et al ²⁴	75 patients with distal radius fractures	Comparison of two external fixators	DASH	Visual analog score for pain
Abramo et al ¹	50 patients with unstable distal radius fractures	Open reduction and internal fixation versus external fixation	DASH	Radiographic assessment

Note. PRWE, Patient-Rated Wrist Evaluation; ROM, range of motion; DASH, Disabilities of the Arm, Shoulder and Hand; SF-36, Short Form 36.

Results

Search Results

The search of Medline, Pubmed, and CINAHL followed by the described 4-stage screening process returned a total of 30 results (see Figure 2).

Patient-Reported Outcome Measures

Nine PROMs were used in these 30 studies. The DASH score was used most frequently, appearing in 20 of the 30 studies. The PRWE was used in eight of the studies, with other measures (Gartland and Werley score, Green and O'Brien score, Michigan Hand Outcomes Questionnaire, Mayo Wrist Score and SF-36) each used on 2 or 3 occasions. The Activities Scale for Kids (ASK) and the Krimmer wrist score were each used on one occasion (Table 2).

Reliability, Validity, and Responsiveness

When assessing reliability, it is important to consider both internal consistency and reproducibility. Internal consistency

is a measure of the correlation of different items in the PROMs that measure a similar parameter. It is generally expressed through examining the correlation of each item with other similar outcome measures to generate a value for Cronbach's alpha. Reproducibility refers to the test-retest properties of a measurement and examines how consistently the same test administrator achieves results for a given test on the same participant in a stable disease state.

Construct validity refers to the degree to which a test measures what it reports to measure, and whether the individual scales behave as expected. Therefore, good construct validity is vital if meaningful clinical conclusions are to be made on the results of a given test. Criterion validity assesses how well a given measurement performs when compared with a "gold-standard" measurement. Responsiveness is the ability of a test to detect important changes in a clinical condition; it is expressed as the standardized response mean (SRM).

A summary of the reliability, validity, and responsiveness findings of the identified PROMs in a hand and wrist trauma population is presented in Table 3. The DASH, PRWE, Michigan Hand Outcomes Questionnaire, SF-36,

Table 2. The PROMs Identified.

PROM	Studies using the PROM	Anatomical region	Assesses	No. of items
DASH	20	Shoulder, arm, hand	Symptoms, function	30
PRWE	8	Wrist/hand	Symptoms, function	15
Gartland and Werley score	4	Wrist/hand	Symptoms, function, deformity	4
Green and O'Brien score	3	Wrist/hand	Symptoms, function, X-ray	5
SF-36	2	Generic	Physical and mental health	36
Michigan Hand Outcomes Questionnaire	2	Hand	Symptoms, function, satisfaction	37
Mayo Wrist Score	2	Wrist	Symptoms, function	4
Krimmer wrist score	1	Wrist	Symptoms, function	4
Activities Scale for Kids	1	General activity	Function	30

Note. PROM, patient-reported outcome measure; DASH, Disabilities of the Arm, Shoulder and Hand; PRWE, Patient-Rated Wrist Evaluation; SF-36, Short Form 36.

Table 3. Summary of Reliability, Validity, and Responsiveness Evidence in Hand and Wrist Trauma Patients.

	Reliability			Responsiveness
	Internal consistency	Reproducibility	Validity (construct)	
DASH	+	+	+	+
PRWE	+	+	+	+
Gartland and Werley	0	0	0	0
Green and O'Brien score	0	0	0	0
Michigan Hand Outcomes Questionnaire	+ (Not trauma)	+ (Not trauma)	+ (Not trauma)	+
Mayo Wrist Score	0	0	0	0
SF-36	+ (Not trauma)	+ (Not trauma)	+	-
Activities Scale for Kids	+ (Not trauma)	+ (Not trauma)	+ (Not trauma)	+ (Not trauma)
Krimmer wrist score	0	0	0	0

Note. + = positive evidence identified; - = negative evidence identified; 0 = no evidence available; (Not trauma) = this addendum was added when the evidence identified had not been reported in a hand and wrist trauma population; DASH, Disabilities of the Arm, Shoulder and Hand; PRWE, Patient-Rated Wrist Evaluation; SF-36, Short Form 36.

and ASK had good evidence of reliability, but this evidence was not gathered in a hand and wrist trauma population for 2 of the measures (SF-36 and ASK). This was also true for validity as the same 5 measures had good evidence of validity, but this had not been assessed in a hand and wrist trauma population for the Michigan Hand Outcomes Questionnaire and ASK. Four measures showed good evidence of responsiveness, with 3 in a hand and wrist trauma population (DASH, PRWE, Michigan Hand Outcomes Questionnaire).

Discussion

The PROM used most often was the DASH questionnaire; the PRWE, Gartland and Werley score, Michigan Hand Outcomes score, Mayo Wrist Score, and SF-36 were also used. Other PROMs such as the Krimmer wrist score and ASK were used just once.

Of the PROMs identified, only the DASH and PRWE have evidence of reliability, validity, and responsiveness in patients with traumatic injuries to the hand and wrist; other measures either have incomplete evidence or evidence gathered in a nontraumatic population.

Disabilities of the Arm, Shoulder and Hand Questionnaire

The DASH questionnaire was developed in 1996, as a joint venture by the American Academy of Orthopaedic Surgeons, the Council of Musculoskeletal Specialty Societies, and the Institute for Work and Health. The aim was to produce a tool capable of evaluating single or multiple disorders specific to the upper limb.²² A 3-stage process with initial item generation by a group of clinical experts and methodologists was used to develop the scale. They subsequently underwent item reduction and field testing.²²

The DASH is a self-administered questionnaire that consists of 30 core questions with an optional 8 further questions, which assess specific work and sport activities.⁴⁸ Each item is scored on a 5-point Likert scale with lower scores indicating with minimal impairment.⁴⁸ Several papers have demonstrated that the DASH is a reliable, valid, and responsive instrument for the assessment of disability and function in patients with upper limb pathology. Beaton et al examined the use of the DASH in 200 patients with a combination of shoulder and wrist pathology and reported good test-retest reliability, validity, and responsiveness in this population.⁴

Lovgren et al looked specifically at the reliability of the DASH in a hand and wrist trauma population and found it to have good internal consistency, with Cronbach's alpha values 0.96 in the Disability subscale and 0.94 in the Optional scale.²⁵ Test-retest reliability was reported as excellent with intraclass correlation coefficient (ICC) of .92. The study also showed construct validity similar to that of the PRWE.²⁵ Dias et al evaluated the validity of the DASH in a population that included trauma patients, finding it to have poor construct validity for nerve disorders but reasonable validity for wrist and finger pathology.¹⁴ The paper also referred to the DASH's "troubling" correlation with established measures, but the comparison was with the Gartland and Werley score for wrist pathology, an outcome measure that has no evidence of reliability, validity, or responsiveness and relies mainly on objective and radiographic measurements.¹⁴ The DASH shows good responsiveness in a hand and wrist trauma population, with a SRM of 2.07 in a population of 59 distal radius patients.²⁸ All of the quoted evidence for the reliability, validity, and responsiveness of the DASH was carried out by teams independent from the designers of the score.

The DASH is well established and was by far the most commonly used PROM found in this study. It has consistently demonstrated good reliability, validity, and responsiveness in several psychometric studies. However, recent work has questioned the validity of the measure in a hand and wrist population.

Patient-Rated Wrist Evaluation

The PRWE was developed in 1998 by MacDermid et al²⁹ in response to a perceived lack of reliable and valid tools for quantifying wrist pain and disability from a patient's perspective. The rating tool was developed using a survey of the members of the International Wrist Investigators Workshop (n = 151). This asked the International Wrist Investigators to define current practice and to take opinions on appropriate content with respect to outcome measurement after wrist injury. A 66% response rate was achieved and the PRWE was developed on the basis of this survey of surgeons. It was subsequently tested for reliability in a

population of hand and wrist trauma patients (64 distal radius fractures and 35 scaphoid fractures); validity was assessed in 101 patients with distal radius fractures.²⁹

Initial reliability testing of the PRWE was undertaken in 3 separate groups. Group 1 had distal radius fractures, group 2 were patients with healed fractures who had undergone physiotherapy, and group 3 were patients 6 years post treatment for scaphoid nonunion.²⁹ This revealed excellent test-retest reliability in all groups with reliability coefficients (ICC) of $>.9$.²⁹ This was reproduced by Lovgren et al, independent from the designers of the test, in 32 patients with distal radius fractures. They reported an ICC $>.9$, which is considered to be excellent.²⁵ They also examined the internal consistency of this PROM with Cronbach's alpha, and this was also high (0.89-0.95).²⁵

MacDermid et al felt that the manner in which the PRWE was designed, using expert surgical opinion, gave it automatic content, or face, validity despite the lack of patient input. Thus, they focused on construct and criterion validity when assessing its validity. Construct validity was demonstrated as statistically significant differences were found in the PRWE over time ($P < .001$) in 101 distal radius fracture patients, which was mirrored by improvement in the SF-36. The amount of improvement was 74% as measured by the PRWE compared with 16% for the SF-36. Criterion validity was measured against the SF-36 and a statistically significant correlation was shown in all domains. This supports the validity of the PRWE in measuring pain and disability in a distal radius fracture population.²⁹ Lovgren et al also examined the validity of the PRWE and reported good convergent validity. They concluded that it was a useful measure in both the acute and rehabilitation phase of distal radius fracture.²⁵

MacDermid et al later demonstrated the responsiveness of the PRWE in 59 distal radius fracture patients, finding that the PRWE, with a SRM of 2.27, was comparable with the DASH (SRM, 2.01). Both the PRWE and the DASH considerably outperformed the SF-36 (SRM, 0.92).²⁷ No data were found on the responsiveness of the PRWE, which was independent of the designers of the score.

Of all the PROMs for the assessment of hand and wrist trauma, the PRWE has by far the best demonstrated reliability and validity in a wrist trauma population. It is the most responsive measure for a distal radius fracture population. It must, however, be remembered that the measure is specific to the wrist and is of little use in assessing a patient with hand injuries.

Gartland and Werley Score

The Gartland and Werley score,¹⁷ originally described in 1951, and subsequently modified by Sarmiento et al in 1975,⁴⁴ is an outcome measure for wrist pathology that combines both subjective and objective factors, recorded by

the observer. Demerit points are awarded for pain, stiffness and disability, range of motion, and distal radio-ulna joint pain with further points relating to complication profile, including radiographic evidence of posttraumatic osteoarthritis.⁵⁰ Sarmiento's modification added points for pronation and grip strength.⁴⁴ The final score is a value between 0 and 52 with low scores corresponding with the best outcome.⁵⁰ Some authors have criticized the scoring system as confusing, because surgeon interpretation can play a significant role, and suggested that using a later modification of the test, by Stewart et al in 1986, would be an improvement.¹³

The measure has been in use since 1951, and in its modified form since 1975, but our study found no evidence to demonstrate its reliability, validity, or responsiveness. Despite the extensive use of this measure in clinical practice, there have been no reliability, validity, or responsiveness studies published to date.

Michigan Hand Outcomes Questionnaire

The Michigan Hand Outcomes Questionnaire, introduced in 1998 by Chung et al, was developed at the Section of Plastic and Reconstructive Surgery at the University of Michigan Medical Centre.¹⁰ The aim was to produce an assessment tool capable of measuring outcomes for patients with all types of hand disorder. After the identification of the key domains for investigation, an initial pool of 100 questions was reduced through pilot testing on 20 patients and follow-up testing on 200 hand surgery patients to a 37-item questionnaire.¹⁰

A 2013 review into the use of the Michigan Hand Outcomes Questionnaire found that it had been used in 58 clinical research studies and 18 studies had investigated the measurement properties of the assessment tool.⁴⁷ Chung et al initially reported the measure to have excellent test-retest reliability with 5 of the 6 scales having an ICC of $>.85$.¹⁰ This has been reproduced in subsequent studies, which found an ICC of .95 in patients with rheumatoid arthritis by a group independent to the Chung's team.³² Chung et al also reported excellent internal consistency with Cronbach's alpha values of 0.84 to 0.93 across the 6 domains.¹⁰ No work has been done to assess these indices in a purely hand and wrist trauma patient group, but Dias et al reported similar internal consistency values of 0.82 to 0.96 in 200 patients that included trauma patients, again independently from the Chung's original design team.

The validity of the Michigan Hand Outcomes Questionnaire was reported by Chung et al who found that construct validity was demonstrated by significant correlation between scales in a predictable pattern. Chung felt that criterion validity was not assessable, given the lack of a gold standard, but the measure compared favorably with the generic Short Form 12 questionnaire.¹⁰ Dias et al also

investigated the validity of the DASH, concluding that there was questionable criterion validity based on poor correlation with the Levine score and Gartland and Werley score; however, it must be remembered that neither of their chosen "gold-standard" outcome measures has proven measurement properties.¹⁴ Responsiveness of the Michigan Hand Outcomes Questionnaire in a sample of 105 patients with hand injuries reported a SRM of 1.05, which was higher than the DASH, but it is not clear whether this has clinical significance.²¹

The Michigan Hand Outcomes Questionnaire has undergone extensive psychometric testing since its introduction in 1998 and has been shown to be a reliable, valid, and responsive tool for outcome assessment in patients with hand injuries. At present, the reliability and validity findings have not been reproduced specifically in a hand trauma population.

Green and O'Brien Score

The Green and O'Brien score, introduced in 1978, is a demerit scoring system assessing pain, disability, range of movement, strength, and radiographic features.¹⁸ The score is primarily evaluated by a doctor but does require a subjective evaluation of pain, which is patient reported.

Despite being one of the first assessment tools available for the assessment of hand and wrist function following injury, its reliability, validity, or responsiveness have never been assessed.

Mayo Wrist Score

Cooney et al modified the Green and O'Brien score in 1987 by changing the demerit items and removing radiographic indices. The resulting outcome measure was referred to as the Mayo Wrist Score.¹² This assesses 4 domains: pain, grip strength, range of motion, and return to employment. Each domain is scored out of 25 points to produce a total score out of 100 points. High scores correlate with good function: Scores of 90 to 100 points denote "excellent" function and a score of less than 65 is considered "poor."

Despite being used by several papers assessing wrist function, no work has been done on the reliability, validity, or responsiveness of the measure.

Short Form 36

The SF-36 is a patient-reported health survey consisting of 8 scaled scores from which final mental and physical scores, each out of 100, are produced. The score was developed as a result of the Medical Outcomes Study undertaken by the Rand Corporation in 1992.⁵³ It is a generic measure of health status and not specific to a particular anatomical region.

The SF-36 has been subject to rigorous psychometric evaluation since its introduction, initially by Ware et al, who reported a reproducibility ICC of $>.8$ for each of the subscales, and $>.9$ for the Physical and Mental Component subscales.^{33,34} The same authors also reported content, construct, and criterion validity in 4 patient groups, which included patients with a combination of medical and psychiatric illnesses.

Our search found little evidence on the reliability of the SF-36 specifically in a hand and wrist trauma patient group. Soohoo et al examined the construct validity and criterion validity of the SF-36 when comparing it with the DASH. They reported a Pearson correlation coefficient of $-.36$ to $-.62$, and noted that the DASH had fewer floor and ceiling scores than the subscales of the SF-36.⁴⁹ MacDermid et al compared the responsiveness of the SF-36 with the DASH and PRWE in a population of distal radius fracture patients and reported that the SF-36 had significantly impaired responsiveness (SRM 0.92) compared with the DASH (SRM 2.01) and PRWE (SRM 2.27).

The SF-36 is a widely used health outcome measure, and has been independently shown to be reliable, valid, and responsive in a generic population. There is limited evidence available regarding its use in a hand and wrist trauma population, and there is evidence to suggest that it is a less valid and responsive measure than other region-specific scoring systems.

Activities Scale for Kids

The ASK is a self-reported questionnaire designed to assess physical function in children. It is not suitable for adults. The scale contains 30 items, in 7 subdomains. These are aggregated into one overall summary score. It is aimed at children aged 5 to 15 years who are suffering from limitation in their physical ability due to musculoskeletal conditions. The scale has been shown to be reliable, valid, and responsive in children with a variety of conditions, including cerebral palsy, juvenile arthritis, and bone tumors. No studies have investigated whether the ASK is reliable, valid, or responsive in a hand and wrist trauma population.

Krimmer Wrist Score

The Krimmer wrist score is a 4-point scoring system. It has been used in the assessment of outcome after wrist fusion, intercarpal fusion, scaphoid nonunion surgery, and other wrist procedures. It assesses power, range of motion, pain, and function to produce a maximum score of 100 points. It is rarely used in the assessment of trauma patients. No evidence of its reliability, validity, or responsiveness for patients with hand or wrist trauma was found.

In conclusion, clinical trials investigating hand and wrist trauma patients have most frequently used the DASH score, which has specific evidence of reliability, validity, and

responsiveness in this population. The DASH is, however, less responsive than the PRWE for patients with distal radius fractures. The PRWE is more site specific than the DASH and also has evidence of reliability, validity, and responsiveness in a hand and wrist trauma population. Despite its excellent psychometric properties, the PRWE is specifically designed for the assessment of wrist pathology, and therefore its use in patients with hand injuries requires further investigation.

The Michigan Hand Outcomes Questionnaire is a widely used and thoroughly investigated assessment tool with evidence of responsiveness in a hand and wrist trauma population, and reliability and validity in a nontrauma population. As suggested in the title, the Michigan Hand Outcomes Questionnaire is specific to the hand and therefore is not intended for use in assessing patients with wrist injuries. Its use in assessing patients with traumatic injuries to the wrist requires further investigation.

The SF-36 and ASK are both generic measures of health status. They both have evidence of reliability, validity, and responsiveness but not in a hand and wrist trauma population. The SF-36 has been shown to be less valid than the DASH, and less responsive than DASH and PRWE in a hand and wrist trauma population.

The Gartland and Werley score, Green and O'Brien score, Mayo Wrist Score, and Krimmer wrist score are well-known assessment tools and are still occasionally used in research. None of these assessment tools has any evidence of reliability, validity, or responsiveness in a hand and wrist trauma population.

Excluding the DASH, and the PRWE, many of the most frequently used PROMs for reporting outcomes in patients with traumatic injuries to the hand and wrist do not have complete evidence of reliability, validity, and responsiveness in a hand and wrist trauma population. This clearly identifies an area where further research is required if these PROMs are to be used in the future to assess patient outcomes following injury and to measure the efficacy of different interventions. It is also worth noting that future validity research into hand and wrist PROMs should be undertaken by comparing the PROM in question, against "gold-standard" measures that have proven measurement properties.

Ethical Approval

Ethical approval for this study was waived by our institutional review board who did not consider it necessary to give approval for a systematic review of the literature on this subject.

Statement of Human and Animal Rights

This article does not contain any studies with human or animal subjects.

Statement of Informed Consent

This study involves no human participants, so informed consent was not sought or required.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Abramo A, Kopylov P, Geijer M, Tägil M. Open reduction and internal fixation compared to closed reduction and external fixation in distal radial fractures: a randomized study of 50 patients. *Acta Orthop*. 2009;80(4):478-485. doi:10.3109/17453670903171875.
- Arora R, Gabl M, Gschwentner M, Deml C, Krappinger D, Lutz M. A comparative study of clinical and radiologic outcomes of unstable colles type distal radius fractures in patients older than 70 years: nonoperative treatment versus volar locking plating. *J Orthop Trauma*. 2009;23(4):237-242. doi:10.1097/BOT.0b013e31819b24e9.
- Atroshi I, Brogren E, Larsson G-U, Kloov J, Hofer M, Berggren A-M. Wrist-bridging versus non-bridging external fixation for displaced distal radius fractures: a randomized assessor-blind clinical trial of 38 patients followed for 1 year. *Acta Orthop*. 2006;77(3):445-453. doi:10.1080/17453670610046389.
- Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C. Measuring the whole or the parts? validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. *J Hand Ther*. 2001;14(2):128-146.
- Belloti JC, Moraes VY, Albers MB, Faloppa F, Santos Dos JBG. Does an ulnar styloid fracture interfere with the results of a distal radius fracture? *J Orthop Sci*. 2010;15(2):216-222. doi:10.1007/s00776-009-1443-7.
- Belloti JC, Tamaoki MJS, Atallah AN, Albertoni WM, Santos dos JBG, Faloppa F. Treatment of reducible unstable fractures of the distal radius in adults: a randomised controlled trial of De Palma percutaneous pinning versus bridging external fixation. *BMC Musculoskelet Disord*. 2010;11:137. doi:10.1186/1471-2474-11-137.
- Boutis K, Willan A, Babyn P, Goeree R. Cast versus splint in children with minimally angulated fractures of the distal radius: a randomized controlled trial. *CMAJ*. 2010;182(14):1507-1512.
- Cassidy C, Jupiter JB, Cohen M, et al. Norian SRS cement compared with conventional fixation in distal radial fractures. A randomized study. *J Bone Joint Surg Am*. 2003;85-A(11):2127-2137.
- Changulani M, Okonkwo U, Keswani T, Kalairajah Y. Outcome evaluation measures for wrist and hand—which one to choose? *Int Orthop*. 2007;32(1):1-6. doi:10.1007/s00264-007-0368-z.
- Chung KC, Pillsbury MS, Walters MR, Hayward RA. Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. *YJHSU*. 1998;23(4):575-587. doi:10.1016/S0363-5023(98)80042-7.
- Cognet JM, Geanah A, Marsal C, Kadoch V, Gouzou S, Simon P. [Plate fixation with locking screw for distal fractures of the radius]. *Rev Chir Orthop Reparatrice Appar Mot*. 2006;92(7):663-672.
- Cooney WP, Bussey R, Dobyns JH, Linscheid RL. Difficult wrist fractures. Perilunate fracture-dislocations of the wrist. *Clin Orthop Relat Res*. 1987;(214):136-147.
- Davis TR. Intra-articular fractures of the distal end of the radius in young adults. *J Bone Joint Surg Am*. 1987;69(5):791.
- 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 hand and the Michigan Hand Outcome Questionnaire. *J Hand Surg Eur Vol*. 2008;33(1):9-17.
- Egol K, Walsh M, Tejwani N, McLaurin T, Wynn C, Paksima N. Bridging external fixation and supplementary Kirschner-wire fixation versus volar locked plating for unstable fractures of the distal radius: a randomised, prospective trial. *J Bone Joint Surg Br*. 2008;90(9):1214-1221. doi:10.1302/0301-620X.90B9.20521.
- Földhazy Z, Törnkvist H, Elmstedt E, Andersson G, Hagsten B, Ahrengart L. Long-term outcome of nonsurgically treated distal radius fractures. *YJHSU*. 2007;32(9):1374-1384. doi:10.1016/j.jhsa.2007.08.019.
- Gartland JJ, Werley CW. Evaluation of healed Colles' fractures. *J Bone Joint Surg Am*. 1951;33-A(4):895-907.
- Green DP, O'Brien ET. Open reduction of carpal dislocations: indications and operative techniques. *YJHSU*. 1978;3(3):250-265.
- Gruber G, Zacherl M, Giessauf C, et al. Quality of life after volar plate fixation of articular fractures of the distal part of the radius. *J Bone Joint Surg Am*. 2010;92(5):1170-1178. doi:10.2106/JBJS.I.00737.
- Hofmeister EP, Kim J, Shin AY. Comparison of 2 methods of immobilization of fifth metacarpal neck fractures: a prospective randomized study. *J Hand Surg Am*. 2008;33(8):1362-1368. doi:10.1016/j.jhsa.2008.04.010.
- Horng Y-S, Lin M-C, Feng C-T, Huang C-H, Wu H-C, Wang J-D. Responsiveness of the Michigan Hand Outcomes Questionnaire and the Disabilities of the Arm, Shoulder, and Hand questionnaire in patients with hand injury. *J Hand Surg Am*. 2010;35(3):430-436. doi:10.1016/j.jhsa.2009.11.016.
- Hudak PL, Amadio PC, Bombardier C, et al. 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*. 1996;29(6):602-608. doi:10.1002/(SICI)1097-0274(199606)29:6<602::AID-AJIM4>3.0.CO;2-L.
- Konrath GA, Bahler S. Open reduction and internal fixation of unstable distal radius fractures: results using the trimed fixation system. *J Orthop Trauma*. 2002;16(8):578-585.
- Krukhaug Y, Ugland S, Lie SA, Hove LM. External fixation of fractures of the distal radius: a randomized comparison of the Hoffman compact II non-bridging fixator and the Dynawrist fixator in 75 patients followed for 1 year. *Acta Orthop*. 2009;80(1):104-108. doi:10.1080/17453670902807433.
- Lövgren A, Hellström K. Reliability and validity of measurement and associations between disability and behavioural factors in patients with Colles' fracture. *Physiother Theory Pract*. 2012;28(3):188-197.
- Lozano-Calderón SA, Souer S, Mudgal C. Wrist mobilization following volar plate fixation of fractures of the distal part of the radius. *J Bone Joint Surg Am*. 2008;90(6):1297-1304.

27. MacDermid JC, Richards RS, Donner A, Bellamy N, Roth JH. Responsiveness of the Short Form-36, Disability of the Arm, Shoulder, and Hand Questionnaire, patient-rated wrist evaluation, and physical impairment measurements in evaluating recovery after a distal radius fracture. *J Hand Surg Am.* 2000;25(2):330-340. doi:10.1053/jhsu.2000.jhsu25a0330.
28. MacDermid JC, Tottenham V. Responsiveness of the Disability of the Arm, Shoulder, and Hand (DASH) and Patient-Rated Wrist/Hand Evaluation (PRWHE) in evaluating change after hand therapy. *J Hand Ther.* 2004;17(1):18-23.
29. MacDermid JC, Turgeon T, Richards RS, Beadle M, Roth JH. Patient rating of wrist pain and disability: a reliable and valid measurement tool. *J Orthop Trauma.* 1998;12(8):577-586.
30. Maciel JS, Taylor NF, McIlveen C. A randomised clinical trial of activity-focussed physiotherapy on patients with distal radius fractures. *Arch Orthop Trauma Surg.* 2005;125(8):515-520. doi:10.1007/s00402-005-0037-x.
31. Magnus CRA, Arnold CM, Johnston G, et al. Cross-education for improving strength and mobility after distal radius fractures: a randomized controlled trial. *Arch Phys Med Rehabil.* 2013;94(7):1247-1255. doi:10.1016/j.apmr.2013.03.005.
32. Massy-Westropp N, Krishnan J, Ahern M. Comparing the AUSCAN Osteoarthritis Hand Index, Michigan Hand Outcomes Questionnaire, and sequential occupational dexterity assessment for patients with rheumatoid arthritis. *J Rheumatol.* 2004;31(10):1996-2001.
33. McHorney CA, Ware JE, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). *Med Care.* 1994 Jan;32(1):40-66.
34. McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care.* 1993;31(3):247-263.
35. McPhail SM, Bagraith KS, Schippers M, Wells PJ, Hatton A. Use of condition-specific patient-reported outcome measures in clinical trials among patients with wrist osteoarthritis: a systematic review. *Adv Orthop.* 2012;2012(2):1-10. doi:10.1186/1471-2474-10-129.
36. Müller M, Chen ZB, Al Morshidy A, Germann G, Sauerbier M. [Postoperative CT-controlled results of renewed osteosynthesis using screw fixation for acute scaphoid fractures]. *Unfallchirurg.* 2006;109(3):185-192. doi:10.1007/s00113-005-1023-z.
37. O'Connor D, Mullett H, Doyle M, Mofidi A, Kutty S, O'Sullivan M. Minimally displaced Colles' fractures: a prospective randomized trial of treatment with a wrist splint or a plaster cast. *J Hand Surg Br.* 2003;28(1):50-53.
38. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care.* 1989;27(3):S217-S232.
39. Pike J, Mulpuri K, Metzger M, Ng G, Wells N, Goetz T. Blinded, prospective, randomized clinical trial comparing volar, dorsal, and custom thermoplastic splinting in treatment of acute mallet finger. *J Hand Surg Am.* 2010;35(4):580-588. doi:10.1016/j.jhsa.2010.01.005.
40. Rhee SH, Lee SK, Lee SL, Kim J, Baek GH, Lee YH. Prospective multicenter trial of modified retrograde percutaneous intramedullary Kirschner wire fixation for displaced metacarpal neck and shaft fractures. *Plast Reconstr Surg.* 2012;129(3):694-703. doi:10.1097/PRS.0b013e3182402e6a.
41. Rostami HR, Arefi A, Tabatabaei S. Effect of mirror therapy on hand function in patients with hand orthopaedic injuries: a randomized controlled trial. *Disabil Rehabil.* 2013;35(19):1647-1651. doi:10.3109/09638288.2012.751132.
42. Rozental TD, Blazar PE, Franko OI, Chacko AT, Earp BE, Day CS. Functional outcomes for unstable distal radial fractures treated with open reduction and internal fixation or closed reduction and percutaneous fixation: a prospective randomized trial. *J Bone Joint Surg Am.* 2009;91(8):1837-1846. doi:10.2106/JBJS.H.01478.
43. Sammer DM, Fuller DS, Kim HM, Chung KC. A comparative study of fragment-specific versus volar plate fixation of distal radius fractures. *Plast Reconstr Surg.* 2008;122(5):1441-1450. doi:10.1097/PRS.0b013e3181891677.
44. Sarmiento A, Pratt GW, Berry NC. Colles' fractures. Functional bracing in supination. *J Bone Joint Surg Am.* 1975;57(3):311-317.
45. Schneiders W, Biewener A, Rammelt S, Rein S, Zwipp H, Amlang M. [Distal radius fracture. Correlation between radiological and functional results]. *Unfallchirurg.* 2006;109(10):837-844. doi:10.1007/s00113-006-1156-8.
46. Schönnemann JO, Hansen TB, Søballe K. Randomised study of non-bridging external fixation compared with intramedullary fixation of unstable distal radial fractures. *J Plast Surg Hand Surg.* 2011;45(4-5):232-237. doi:10.3109/2000656X.2011.613243.
47. Shauver MJ, Chung KC. The Michigan Hand Outcomes Questionnaire after 15 years of field trial. *Plast Reconstr Surg.* 2013;131(5):779e-787e. doi:10.1097/PRS.0b013e3182865d83.
48. Smith MV. Upper extremity-specific measures of disability and outcomes in orthopaedic surgery. *J Bone Joint Surg Am.* 2012;94(3):277-285. doi:10.2106/JBJS.J.01744.
49. SooHoo NF, McDonald AP, Seiler JG, McGillivray GR. Evaluation of the construct validity of the DASH questionnaire by correlation to the SF-36. *YJHSU.* 2002;27(3):537-541.
50. Souer J-S, Lozano-Calderon SA, Ring D. Predictors of wrist function and health status after operative treatment of fractures of the distal radius. *YJHSU.* 2008;33(2):157-163. doi:10.1016/j.jhsa.2007.10.003.
51. Varitimidis SE, Basdekis GK. Treatment of intra-articular fractures of the distal radius fluoroscopic or arthroscopic reduction? *J Bone Joint Surg Br.* 2008;90(6):778-785.
52. Vinnars B, Pietreanu M, Bodstedt A, Ekenstam FA, Gerdin B. Nonoperative compared with operative treatment of acute scaphoid fractures: a randomized clinical trial. *J Bone Joint Surg Am.* 2008;90(6):1176-1185. doi:10.2106/JBJS.G.00673.
53. Ware JE Jr, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): I. Conceptual framework and item selection. *Med Care.* 1992;30(6):473-483.
54. Wei DH, Raizman NM, Bottino CJ, Jobin CM. Unstable distal radial fractures treated with external fixation, a radial column plate, or a volar plate: a prospective randomized trial. *J Bone Joint Surg Am.* 2009;91(7):1568-1577.
55. Wells GA, Russell AS, Haraoui B, Bissonnette R, Ware CF. Validity of quality of life measurement tools—from generic to disease-specific. *J Rheumatol Suppl.* 2011;88:2-6. doi:10.3899/jrheum.110906.
56. Wilcke MKT, Abbaszadegan H, Adolphson PY. Wrist function recovers more rapidly after volar locked plating than after external fixation but the outcomes are similar after 1 year. *Acta Orthop.* 2011;82(1):76-81. doi:10.3109/17453674.2011.552781.