



Contents lists available at ScienceDirect

Journal of Hand Surgery Global Online

journal homepage: www.JHSGO.org

Original Research

Measuring Outcomes Over Time in Distal Radius Fractures: A Comparison of Generic, Upper Extremity-Specific and Wrist-Specific Outcome Measures



Amarpal S. Cheema, MD, * Peter J. O'Brien, MD, * Henry M. Broekhuyse, MD, * Kelly A. Lefavre, MD *

* Department of Orthopaedic Surgery, Division of Orthopaedic Trauma, University of British Columbia, Vancouver, British Columbia, Canada

ARTICLE INFO

Article history:

Received for publication May 1, 2020
 Accepted in revised form June 25, 2020
 Available online August 14, 2020

Key words:

Disabilities of the Arm, Shoulder, and Hand
 Distal radius
 Fracture
 Functional outcome scores
 Patient-Rated Wrist Evaluation

Purpose: This study compared the responsiveness of a generic (Short Form-36 [SF-36]), an upper extremity-specific (Disabilities of the Arm, Shoulder, and Hand [DASH]) and a wrist-specific (Patient-Rated Wrist Evaluation [PRWE]) outcome score when evaluating distal radius fractures over time.

Methods: We observed 235 patients who met the inclusion criteria of an isolated distal radius fracture treated surgically or nonsurgically and greater than age 50 years for 12 months in this prospective study. Standardized assessments were performed at baseline and at 6 and 12 months. Exclusion criteria included subjects with concomitant injuries in the ipsilateral limb and follow-up of less than 1 year. Responsiveness was evaluated through the standardized response mean and the proportion who met a minimal clinically important difference. Floor and ceiling effects were also calculated.

Results: The standardized response mean was significantly greatest for the DASH between baseline and 6 months ($P < .001$), and the PRWE between both baseline and 6 months ($P < .01$) and 6 and 12 months ($P < .01$) compared with the SF-36. The proportion of patients who met a minimal clinically important difference between baseline and 6 months was greater in the PRWE, but it did not meet statistical significance ($P = .12$). The PRWE demonstrated a high ceiling effect at baseline (76.6%) but less so at 12 months (16.9%). The DASH demonstrated similar ceiling effects at baseline (62.9%) and 12 months (18.6%). The SF-36 had no ceiling effect.

Conclusions: In the first 6 months, both the DASH and PRWE have greater responsiveness in assessing change over the SF-36 in distal radius fractures. From 6 to 12 months, the wrist-specific PRWE has greater responsiveness over both the DASH and SF-36. This supports the use of the anatomy- and injury-specific outcome measures over the generic outcome measure in detecting change over a patient's early recovery. However, as the time from injury increases, the absence of a ceiling effect from the generic outcome measure may become more useful.

Clinical relevance: This study demonstrates the responsiveness of the DASH, PRWE, and SF36 in assessing distal radius fractures treated in patients greater than age 50 in the first year. In establishing the most responsive measure, respondent burden can be decreased in future research.

Crown Copyright © 2020, Published by Elsevier Inc. on behalf of The American Society for Surgery of the Hand. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Declaration of interests: No benefits in any form have been received or will be received by the authors related directly or indirectly to the subject of this article. No funding was received for this article.

Corresponding author: Kelly A. Lefavre, MD, MSc, FRCSC, Department of Orthopaedics, Division of Orthopaedic Trauma, University of British Columbia, VGH Research Pavilion, 110-828 West 10th Avenue, Vancouver, BC, Canada V5Z 1M9.

E-mail address: kelly.lefavre@vch.ca (K.A. Lefavre).

<https://doi.org/10.1016/j.jhsg.2020.06.006>

2589-5141/Crown Copyright © 2020, Published by Elsevier Inc. on behalf of The American Society for Surgery of the Hand. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Distal radius fractures are the most common upper-extremity fracture and the second most common overall fracture.¹ The annual incidence in the United States is estimated to be 643,000 yearly.¹ They represent a diverse spectrum of injury with numerous treatment strategies. To date, radiological parameters and objective measures such as range of motion and grip strength have been outcome surrogates.²

However, the current standard in orthopedic care is the use of functional outcome scores that are patient-rated.^{3,4} The most

widely used is the Short Form-36 (SF36), a validated generic functional outcome measure.⁵ Specifically in orthopedics, the Physical Component Score (PCS) is used.⁵

Following the success of generic outcome measures, injury- and anatomy-specific outcome measures were developed in the hope that they would be more sensitive to changing functional status. In measuring treatment outcomes for distal radius fractures, the most commonly used outcome measures have been the wrist-specific, Patient-Rated Wrist Evaluation (PRWE) and the upper extremity-specific, Disabilities of the Arm, Shoulder, and Hand (DASH). Both of these have established validity and reliability.^{6–8}

Although in principle, injury- or anatomy-specific patient-reported outcome measures (PROMs) are more sensitive to change, this is not the case in the lower extremities. Studies evaluating tibial shaft fractures, tibial plateau fractures, and pilon fractures have shown superior responsiveness of the SF-36 to musculoskeletal and anatomy-specific measures.^{9–11}

However, there is a scarcity of literature identifying the responsiveness over time of these different outcome measures in upper-extremity injuries. Once validation and reliability are established, responsiveness becomes a distinguishing factor for outcome measure utility because it is best able to track a patient's recovery. The greater the responsiveness of an outcome measure, the more sensitive it is to detecting change over time.¹² To establish responsiveness, outcome measurement scores at multiple time points are required.^{13,14}

Amadio et al¹⁵ compared responsiveness between the SF-36, a modified arthritis impact measurement scale (AIMS2), and a modified Brigham and Women's carpal tunnel instrument. A total of 21 patients were involved, all of whom had only Colles fractures. The single interval was between the removal of immobilization and 3 months. The authors found the highest responsiveness for function-related subscales of their measures. Unfortunately, the AIMS2 and modified Brigham and Women's carpal tunnel instrument are not frequently used in the distal radius fracture literature. Previous studies looking at the responsiveness of generic and upper extremity-specific instruments in this population have been limited by sample size, length of follow-up, and incompleteness of methodology.

The objective of this study was to evaluate the responsiveness of the SF-36, the DASH, and the PRWE in distal radius fracture patients over the first year of recovery. This will help establish the most appropriate outcome measure to use in the setting of distal radius fracture research. The hypothesis was that the upper-extremity and wrist-specific measures would demonstrate greater responsiveness than the generic outcome measure.

Materials and Methods

A prospective study was conducted at a single Level 1 trauma center between 2007 and 2011 involving patients who received treatment for a distal radius fracture. All patients greater than age 50 years who sustained an isolated distal radius fracture and were receiving surgical and nonsurgical treatment were approached for study enrollment. This study was approved by the institutional ethics committee at our institution. Informed consent was obtained from all patients before enrollment. Exclusion criteria included subjects with concomitant injuries in the ipsilateral limb and those with follow-up of less than 1 year. We collected basic demographic and injury information.

Three PROMs were used in this study: the SF-36, DASH, and PRWE. The SF-36 is a generic health measure evaluating 8 domains: physical functioning, role limitations owing to physical health, bodily pain, general health perceptions, vitality, social functioning, role limitations owing to emotional problems, and general mental

health.¹⁶ Raw scores are scaled from 0 to 100, in which in each case a lower score indicates more disability. These scaled scores are standardized to z-scores (based on a US population), linearly transformed to a mean of 50 and an SD of 10, and then combined using specified factor score coefficients into 2 summary measures, the PCS and the Mental Component Score.¹⁶ The PCS was used in this study.

The DASH is widely used in the orthopedic literature for upper-extremity injuries. It is a 30-item questionnaire in which each item is scored on a 5-point scale.⁸ These scores are then scaled from 0 to 100, in which a greater score represents greater disability. Because a lower score on the SF-36 represents greater disability, the DASH score is inverted to allow for easier comparison. Reliability and validity of the DASH have been evaluated in patients with disorders of the shoulder, elbow, wrist, and hand.⁸ This was selected over the QuickDASH because it uses 30 items as opposed to 11, and full-length versions of the generic measure (SF-36) were used.³

The PRWE consists of 2 sections that investigate pain with 5 items and function with 10.⁸ Total scores add to a maximum of 100 and do not require scaling. No inversion of scores is required because a greater score corresponds to less disability. This score was originally designed to assess outcomes specifically in distal radius fractures.^{8,17}

All patients who completed questionnaires at 3 time points were included in the responsiveness analysis. Patients were asked to determine the preinjury function score (baseline) within 2 weeks of the injury, and at 6 months and 1 year after the injury. Pearson correlation coefficients were calculated between all pairs of scores.

For consistent comparison of change between time points, only subjects with complete data at all time points were included in the final analysis. Of the 236 patients, 124 (53%) completed all 3 questionnaires, SF-36, DASH, and PRWE, for all time points (Table 1) and were included in the final analysis. For patients with complete data, 107 (86.3%) were female, mean age 64 years; they had a mean Injury Severity Score of 9.0, confirming the isolated nature of the injuries. There was no statistically significant difference in mean age, sex, or mean Injury Severity Score between patients with complete or incomplete data. To ensure the following results are based on the same cohort of patients, they are limited to the 124 patients with complete data.

The primary outcome measure to compare responsiveness was the standardized response mean (SRM); secondary comparisons were the proportion of patients achieving the minimal clinically important difference (MCID), and the floor and ceiling effects. The SRM is the mean score improvement divided by the SD of score improvement.^{13,14} An SRM was calculated between baseline and 6 months, and 6 months and 1 year for all scores. The SRMs for different scores over the same period were compared using the paired *t* test on patient-specific standardized score improvements. The MCID is the smallest change in score that reflects a clinically notable difference and is specific to the disease. There is no established consensus in the literature regarding the MCID for these outcome measures in patients treated with distal radius fractures. As such, an accepted method for estimating MCID is to use one-half the SD of the patient scores when they are maximally affected by the disease.¹⁸ Therefore, the MCID was calculated based on data at the 6-month point using the statistical method described by Norman et al.¹⁸ The proportions of patients who met an MCID on different scores over the same period were compared using the McNemar test. To describe ceiling and floor effects for the measures, the proportions of patients who achieved the maximum and minimum values detectable by each outcome score are reported for each time point.

All statistical analyses were carried out using the R statistical computing environment (R Core Team, R Foundation for Statistical

Table 1
Demographic Data of Study Patients Comparing Groups Based on Completeness of Data

Characteristics	Patients With Complete Data	Patients With Incomplete Data	P Value
n (%)	124 (52.5)	112 (47.5)	
Sex, n (%)			
Male	17 (13.7)	21 (19.1)	1*
Female	107 (86.3)	89 (80.9)	1*
Age, y			
Mean (SD)	64 (7)	65 (8)	0.35†
Median (range)	63 (55–95)	64 (52–86)	
Injury Severity Score			
Mean (SD)	9.0 (0.8)	9.0 (1.6)	0.94†
Median (range)	9.0 (4–14)	9.0 (4–22)	
>9 (%)	1 (0.8)	2 (1.8)	1*
>18 (%)	0	1 (0.9)	1*

* Fisher exact test.

† Student t test.

Computing, Vienna, Austria, 2018). $P < .05$ was considered statistically significant.

Results

Figure 1 shows the distribution of scores for the SF-36-PCS, PRWE, and DASH at each time point. The scores correlated well with each other; all pairwise correlations between time points were highly statistically significant at $P < .001$ (Table 2).

The SRM was calculated between baseline and 6 months, and 6 and 12 months (Fig. 2). The SRMs of the PRWE and DASH were significantly greater in magnitude than the SF-36 between baseline and 6 months ($P < .01$ and $P < .001$, respectively). Between 6 and 12 months, the SRM of the PRWE was significantly greater in magnitude than the SF-36 ($P < .01$) and the DASH ($P < .04$). There were no differences in the SRM between the PRWE and DASH from baseline to 6 months.

The MCID was calculated for each score between time points using one-half of an SD of the outcome scores at 6 months, when the patient was most affected by the injury (Fig. 3). Using this method, the MCID for the SF-36 PCS was 4.09, the PRWE was 8.47, and the DASH was 6.82. There was no statistical difference in the proportion of patients who reached MCID among the SF-36, PRWE, and DASH.

With respect to floor effects, no patients achieved the lowest level of functioning for the SF-36 or DASH at any point. For the PRWE, 2 (1.0%), 18 (9.0%), and 12 (6.4%) patients achieved the lowest level of function at baseline, 6 months and 12 months, respectively. A clinically acceptable floor effect is 10%.¹⁰

With regard to ceiling effects (Table 3), no patients achieved the highest level of functioning that could be assessed by the SF-36 at any time point, whereas the PRWE and DASH exhibited significant ceiling effect at all time points. At baseline, the PRWE and DASH yielded 74.9% and 61.4% of patients achieving the highest level of functioning measurable, which decreased to 8.5% and 8.1% at 6 months and 17.5% and 16.4% at 12 months. A clinically acceptable ceiling effect is 10%.¹⁰

Discussion

Assessing treatment success using functional outcome measures has become standard practice in orthopedics.^{3,4,19} They are patient-centered and therefore considered superior to using outcome surrogates such as radiographic parameters.^{3,4,19} Generic outcome scores have been useful because they can translate across differing disease and injury states, are widely recognized, and have the most literature testing their utility.²⁰ With orthopedic injuries,

it has been posited that anatomy- or injury-specific outcome measures will ascertain outcome more accurately.²¹ However, in recent years, previous psychometric analyses regarding outcomes after tibial shaft, plateau, and pilon fractures have demonstrated no advantages from using a disease-specific score over the generic SF-36 score in terms of responsiveness.^{9–11}

It is important to establish the most responsive outcome measure in a given patient and injury population if it is to be used.²² Doing so reduces the respondent burden put on study participants, because the less-responsive measures can be foregone. It is established that lengthy questionnaires result in decreased compliance.²² In patient-reported outcomes research, minimizing respondent burden is considered fundamental to operationalizing measurement tools.²¹ This allows for streamlined research, yielding data of most interest and avoiding unnecessary statistical analyses. It maximizes study participation and follow-up, thereby increasing study power.²²

Jayakumar et al³ performed a systematic review of outcome measures in upper-extremity trauma and found 144 total measures, with 7 applicable to the arm in general, 59 specific to the wrist, and 20% overall related to trauma. Kleinlugtenbelt et al⁶ examined the validity of outcome measures used specifically for distal radius fractures and found 12. It is self-evident that if a study participant is given such a plethora of outcome measures to complete, the response rate will diminish. Even when administering only the most common, DASH and PRWE in addition to the widely used SF-36, it will certainly yield a lower response rate than a sole questionnaire that is established as the most suitable for the given disease entity. In that regard, the most suitable outcome measure should be the one that best detects change over time. Good evidence is lacking on what that measure is in the distal radius fracture population.

In this study, the upper extremity-specific DASH and wrist-specific PRWE demonstrated the greatest ability to detect change over time compared with the generic SF-36. Both the PRWE and DASH yielded the greater SRM for each time point, indicating that they are more responsive to detecting change. Between baseline and 6 months, both the PRWE and DASH demonstrated statistically greater responsiveness than the SF-36, with the greatest magnitude SRM belonging to the DASH. From 6 to 12 months, only the PRWE demonstrated statistically greater responsiveness over the SF-36 and DASH. This provides evidence that in the setting of distal radius fractures, the most responsive outcome measure to track change over the first year is the PRWE. The PRWE originated to measure outcomes specifically for distal radius fractures and is therefore fitting as the most responsive score for such a setting.¹⁷ This expands on

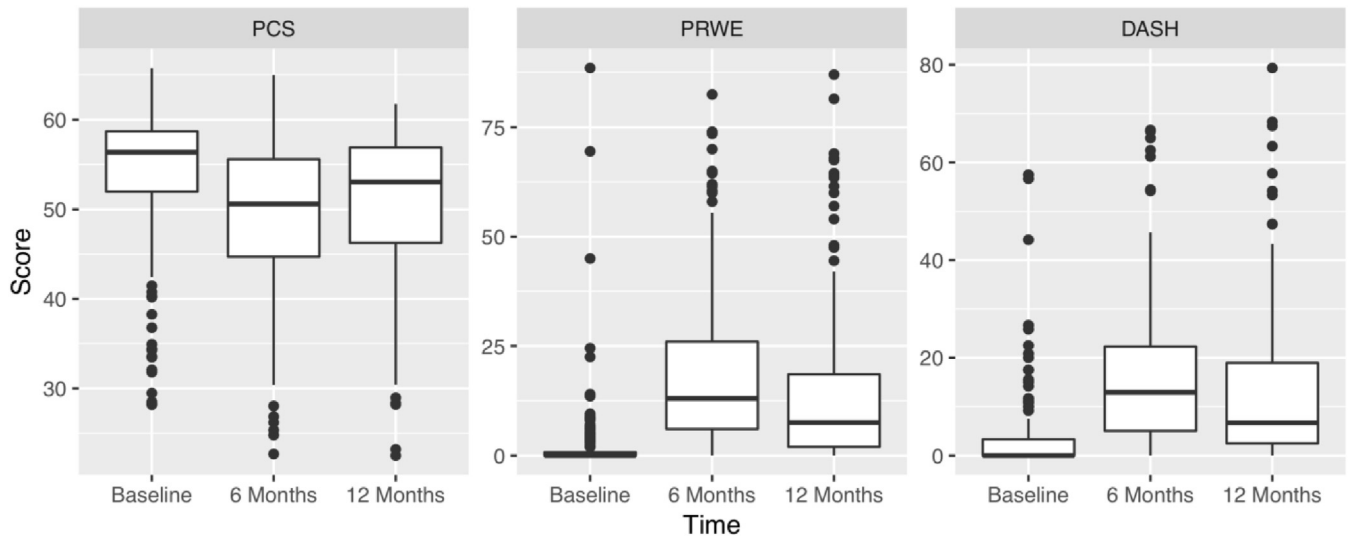


Figure 1. Distribution of scores at each time point.

Table 2
Pearson Correlations Among SF-36 PCS, PRWE, and DASH for Patients With Complete Data

Measures Compared	Time Point		
	Baseline	6 Mo	12 Mo
SF-36 and PRWE	−0.28	−0.51	−0.55
SF-36 and DASH	−0.56	−0.69	−0.67
PRWE and DASH	0.44	0.77	0.84

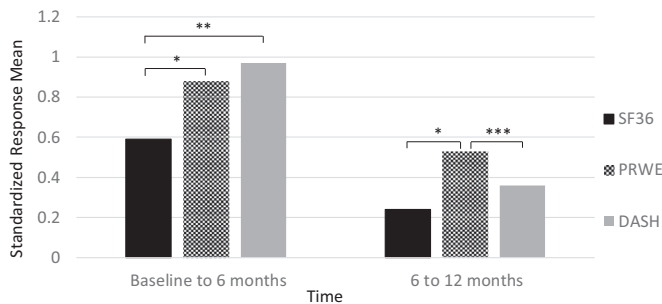


Figure 2. Comparison of magnitude of SRM for SF-36, PRWE, and DASH. ¹The mean improvements are all negative from baseline to 6 months and positive from 6 months to 12 months. * P<0.01; **P<0.001; ***P<0.04.

previous work by MacDermid et al,²³ who demonstrated that both the PRWE and DASH were more responsive than the SF-36, and that the PRWE yielded greater SRMs than the DASH over the early stages after treatment. In that study, the responsiveness of the DASH was greatest between baseline and 3 months and decreased at the 6-month follow-up. The current study revealed the continuing trajectory of the DASH responsiveness, which equalizes with the SF-36 by 1 year. This study also confirms that although there are many other health drivers in an older patient population, the upper-extremity scores remain superior in detecting change.

Although the upper-extremity and wrist-specific scores have shown greater responsiveness, they demonstrate ceiling effect at the beginning and end of treatment.²⁴ MacDermid et al²³ did not investigate ceiling effects, nor did they investigate floor effects. The

current study investigated floor and ceiling effects, and the SF-36 was found to have neither at any time points measured. For both the DASH and PRWE, greater than 60% of patients at baseline achieved the maximal functional level. Although this decreased after injury, it was still elevated at over 8% at 6 months and over 16% at 12 months. Thus, greater follow-up would demonstrate an increasing ceiling effect from the non-generic measures. In other words, as a patient’s function improves, the non-generic measures carry greater ability to detect that change. However, as they plateau, they fail to distinguish among patients who are high-functioning.

There were limitations to this study. A total of 47% of patients who were enrolled in the study did not complete all outcome scores at all time points, which left only 53% for full analysis. Fortunately, we found no differences between patients with incomplete data and those with complete data. The loss of full follow-up was unsurprising given the traumatic nature of these injuries. This study also did not address change beyond 12 months and therefore cannot establish when a patient’s function plateaus. Third, the patient cohort underwent both nonsurgical and surgical care, which may have altered responsiveness. Fourth, the preinjury baseline outcome scores in operatively treated patients were obtained after surgery before discharge from the hospital, which may have resulted in recall bias. However, in the setting of elective total hip arthroplasty, patients were found to be accurate when using the SF-36 to recall the baseline level of function.²⁵ Finally, newer methods of assessing patient-reported outcome are being developed that were not investigated. The second version of the Patient-Reported Outcomes Measurement Information System Upper Extremity aims to standardize and streamline the collection and interpretation of PROMs. This is used in combination with computerized adaptive testing and allows for question selection based on previous answers, creating a tailored questionnaire that increases accuracy and efficiency.²⁶

For distal radius fractures in older adults, the upper extremity-specific DASH and wrist-specific PRWE demonstrate the greatest ability to detect change over time compared with the generic SF-36 in the first 6 months, whereas the PRWE continues to demonstrate better responsiveness for up to a year.

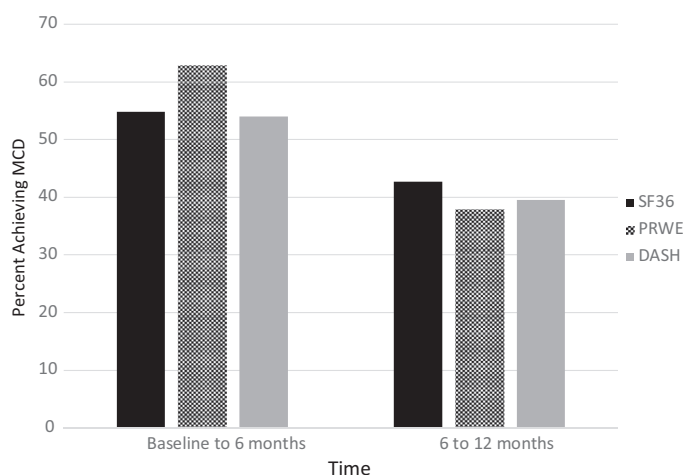


Figure 3. Percentage of patients achieving MCID between time points for SF-36, PRWE, and DASH.

Table 3
Ceiling Effects at All Time Points: Patients at Highest Possible Level of Functioning (n [%])

Outcome Measures	Time Point		
	Baseline	6 Mo	12 Mo
SF-36	0	0	0
PRWE	95 (76.61)	10 (8.06)	21 (16.94)
DASH	78 (62.90)	9 (7.26)	23 (18.55)

References

- Mosenthal WP, Boyajian HH, Ham SA, Conti Mica MA. Treatment trends, complications, and effects of comorbidities on distal radius fractures. *Hand (N Y)*. 2019;14(4):534–539.
- Goldhahn J, Beaton D, Ladd A, Macdermid J, Hoang-Kim A. Recommendation for measuring clinical outcome in distal radius fractures: a core set of domains for standardized reporting in clinical practice and research. *Arch Orthop Trauma Surg*. 2014;134(2):197–205.
- Jayakumar P, Williams M, Ring D, Lamb S, Gwilym S. A systematic review of outcome measures assessing disability following upper extremity trauma. *JAAOS Glob Res Rev*. 2017;1(4):e021.
- Swiontkowski ME. The outcomes movement in orthopaedic surgery: where we are and where we should go. *J Bone Joint Surg Am*. 1999;81(5):732–740.
- Laucis NC, Hays RD, Bhattacharyya T. Scoring the SF-36 in orthopaedics: a brief guide. *J Bone Joint Surg Am*. 2014;97(19):1628–1634.
- Kleinlugtenbelt YV, Nienhuis RW, Bhandari M, Goslings JC, Poolman RW, Scholtes VAB. Are validated outcome measures used in distal radial fractures truly valid? A critical assessment using the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) checklist. *Bone Joint Res*. 2016;5(4):153–161.
- Ritting AW, Wolf JM. How to measure outcomes of distal radius fracture treatment. *Hand Clin*. 2012;28(2):165–175.
- Dowrick AS, Gabbe BJ, Williamson OD, Cameron PA. Outcome instruments for the assessment of the upper extremity following trauma: a review. *Injury*. 2005;36(4):468–476.
- Busse JW, Bhandari M, Guyatt GH, et al. Use of both short musculoskeletal function assessment questionnaire and short form-36 among tibial-fracture patients was redundant. *J Clin Epidemiol*. 2009;62(11):1210–1217.
- Dattani R, Slobogean GP, O'Brien PJ, et al. Psychometric analysis of measuring functional outcomes in tibial plateau fractures using the Short Form 36 (SF-36), Short Musculoskeletal Function Assessment (SMFA) and the Western Ontario McMaster Osteoarthritis (WOMAC) questionnaires. *Injury*. 2013;44(6):825–829.
- Sepehri A, Lefavre KA, Brien PJO, Broekhuysen HM, Guy P. Comparison of generic, musculoskeletal-specific, and foot and ankle-specific outcome measures over time in tibial plafond fractures. *Foot Ankle Orthop*. 2019;4(4):1–7.
- Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chronic Dis*. 1985;38(1):27–36.
- Stratford PW, Binkley JM, Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. *Phys Ther*. 1996;76(10):1109–1123.
- Liang MH, Lew RA, Stucki G, Fortin PR, Daltroy L. Measuring clinically important changes with patient-oriented questionnaires. *Med Care*. 2002;40(4):II45–II51.
- Amadio PC, Silverstein MD, Ilstrup DM, Schleck CD, Jensen LM. Outcome after Colles fracture: the relative responsiveness of three questionnaires and physical examination measures. *J Hand Surg Am*. 1996;21(5):781–787.
- Ware JE, 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.
- MacDermid JC. Patient rating of wrist pain and disability: a reliable and valid measurement tool. *J Orthop Trauma*. 1998;12(8):577–586.
- Norman GR, Sloan JA, Wywich KW. Interpretation of changes in health-related quality of life the remarkable universality of half a standard deviation. *Med Care*. 2003;41(5):582–592.
- Ayers DC, Bozic KJ. The importance of outcome measurement in orthopaedics. *Clin Orthop Relat Res*. 2013;471(11):3409–3411.
- Beaton DE, Schemitsch E. Measures of health-related quality of life and physical function. *Clin Orthop Relat Res*. 2003;413:90–105.
- Changulani M, Okonkwo U, Keswani T, Kalairajah Y. Outcome evaluation measures for wrist and hand—Which one to choose? *Int Orthop*. 2008;32(1):1–6.
- Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM. Developing a Protocol for Observational Comparative Effectiveness Research: A User Guide. Rockville (MD): Agency for Healthcare Research and Quality (US). 2013.
- 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.
- Kim SJ, Lee BG, Lee CH, Choi WS, Kim JH, Lee KH. Comparison of ceiling effects between two patient-rating scores and a physician-rating score in the assessment of outcome after the surgical treatment of distal radial fractures. *Bone Joint J*. 2015;97-B(12):1651–1656.
- Marsh J, Bryant D, MacDonald SJ. Older patients can accurately recall their preoperative health status six weeks following total hip arthroplasty. *J Bone Joint Surg A*. 2009;91(12):2827–2837.
- Tyser AR, Hung M, Bounsanga J, Voss MW, Kazmers NH. Evaluation of version 2.0 of the PROMIS Upper Extremity computer adaptive test in non-shoulder upper extremity patients. *J Hand Surg Am*. 2019;44(4):267–273.