TOPIC HIGHLIGHT

WJO 5th Anniversary Special Issues (7): Shoulder

Functional outcomes assessment in shoulder surgery

James D Wylie, James T Beckmann, Erin Granger, Robert Z Tashjian

James D Wylie, James T Beckmann, Erin Granger, Robert Z Tashjian, Department of Orthopaedic Surgery, University of Utah, Salt Lake City, UT 84108, United States
Author contributions: Wylie JD, Beckmann JT, Granger E and Tashjian RZ solely contributed to the concept, writing and editing of this paper.
Correspondence to: Robert Z Tashjian, MD, Department of Orthopaedic Surgery University of Utah, 590 Wakara Way, Salt Lake City, UT 84108,
United States. robert.tashjian@hsc.utah.edu
Telephone: +1-801-5875457 Fax: +1-801-5875411
Received: May 26, 2014   Revised: July 3, 2014
Accepted: July 18, 2014
Published online: November 18, 2014

Abstract

The effective evaluation and management of orthopaedic conditions including shoulder disorders relies upon understanding the level of disability created by the disease process. Validated outcome measures are critical to the evaluation process. Traditionally, outcome measures have been physician derived objective evaluations including range of motion and radiologic evaluations. However, these measures can marginalize a patient's perception of their disability or outcome. As a result of these limitations, patient self-reported outcomes measures have become popular over the last quarter century and are currently primary tools to evaluate outcomes of treatment. Patient reported outcomes measures can be general health related quality of life measures, health utility measures, region specific health related quality of life measures or condition specific measures. Several patients self-reported outcomes measures have been developed and validated for evaluating patients with shoulder disorders. Computer adaptive testing will likely play an important role in the arsenal of measures used to evaluate shoulder patients in the future. The purpose of this article is to review the general health related quality-of-life measures as well as the joint-specific and condition specific measures utilized in evaluating patients with shoulder conditions. Advances in computer adaptive testing as it relates to assessing dysfunction in shoulder conditions will also be reviewed.

Core tip: Health related quality of life evaluation includes general health measures, health utility measures, general shoulder measures and condition specific shoulder measures. A combination of a general/health utility measure with a shoulder measure or condition specific measure is needed to fully capture outcomes in the treatment of shoulder conditions.

Wylie JD, Beckmann JT, Granger E, Tashjian RZ. Functional outcomes assessment in shoulder surgery. *World J Orthop* 2014; 5(5): 623-633 Available from: URL: http://www.wjgnet.com/2218-5836/full/v5/i5/623.htm DOI: http://dx.doi.org/10.5312/wjo.v5.i5.623

INTRODUCTION

Measuring outcome of orthopedic procedures has changed remarkably over the last twenty to thirty years. Objective physician measurements in large part have given way to subjective patient reported outcome measures[1]. The driving force for this was the inherent bias in the clinician assessment along with how this assessment method marginalized the patient's perception of their outcome[2-4]. Quality of life is the main outcome measure in orthopedics due to the simple fact that most orthopaedic interventions do not increase a patient's life span, so survival is not a realistic outcome measure. A growing body of literature has evolved over the past 30 years regarding measurement of health related quality of life in orthopedic patients, and more specifically patients with

shoulder disorders, both at baseline and after operative or non-operative intervention.

Patient reported health related quality of life (HRQoL) can be measured in multiple ways. Activity level measures reflect the effect of a disease or intervention on a patient's ability to recreate; these are more commonly used in the evaluation of lower extremity disease because the lower extremity has a more profound effect upon patient activity[5]. However, a shoulder activity level questionnaire has been published but so far used sparingly in the literature[6,7]. General HRQoL measures evaluate the effect of a condition on the patients overall health. These measures may be less responsive to shoulder diseases and their treatment because they are designed to evaluate patients general well being[8]. Health utility measures allow calculation of quality adjusted life years (QALY) and are used to economically evaluate treatments[9]. Shoulder HRQoL measures are designed to either be general shoulder measures or condition specific measures that are validated only for certain diagnoses[10,11]. The more specific the measure, the more change the can be elucidated in treatment of a given shoulder condition, however the more general the measure the better it judges the patient's change in overall health. This article will review the various patient reported measures used to evaluate patients with shoulder disorders and the outcome of their treatment.

GENERAL HRQOL AND HEALTH UTILITY MEASURES

General HRQoL measures are commonly used across medical specialties. This fact underscores their importance in the evaluation of patient outcomes from orthopedic conditions. Including general health measures in the evaluation of shoulder treatment allows the comparison of quality of life improvements from orthopedic intervention to those across other organ systems, and can be used to compare the effect of treatment of shoulder conditions to those of the lower extremities. In the current era of healthcare reform, using general HRQoL and health utility measures to evaluate the effect of shoulder treatment compared to that of other conditions across the body will be important to justify health care dollars for the treatment of shoulder disease[12]. This section will review the most commonly used general HRQoL and health utility measures in orthopaedics.

The medical outcomes study short form-36 (SF-36) was originally described in 1992 and is the most commonly used tool to assess general health-related quality of life[13,14]. The Sickness Impact Profile and the Nottingham Health Profile are other general health related quality of life measures that are less commonly used in the orthopaedic literature but their properties and usefulness in musculoskeletal conditions have been recently reviewed[15]. The SF-36 measures eight dimensions of general health [physical functioning (PF), Role Physical, Bodily Pain, General Health, Vitality, Social Functioning (SF), Role Emotional, and Mental Health] and has two summary scores [physical component score (PCS) and mental -component score (MCS)]. The scales are scored based on general United States population norms with a mean of 50 and a standard deviation of 10, with a score of 50 for each scale representing "average" health. The Short-Form-12 (SF-12) was described shortly thereafter and is a validated brief subset of the SF-36 that provides good approximation of the SF-36 summary scores (PCS and MCS) and only moderate approximation of the eight SF-36 domains[16].

The Short Forms are the most commonly used general health related quality of life measure in orthopedics and medical research at large[8,17]. It has been extensively validated and is responsive to treatment of many disease states. For detailed explanation, extensive references, and information on use of the Short Forms please visit www.sf-36.org. The SF-36 has been used to evaluate patients after a variety of shoulder surgical procedures including rotator cuff repair, anatomic shoulder arthroplasty and reverse total shoulder arthroplasty[18,19]. In general the PCS component of the SF-36 score improves after surgical treatment while the MCS component has little change[19,20]. Rotator cuff tears, glenohumeral arthritis, anterior glenohumeral instability, adhesive capsulitis and impingement have been determined to rank in severity with hypertension, congestive heart failure, acute myocardial infarction, diabetes mellitus and clinical depression as evaluated by the SF-36[21]. Boorman et al[22] has found that while anatomic total shoulder arthroplasty does not restore general health status to age adjusted controls it does provide improvement to the same level as seen after coronary artery bypass surgery. Finally, most shoulder outcome instruments do not adequately reflect general health-related quality of life[23,24]. Consequently, inclusion of general health-related quality of life measures in the evaluation of shoulder conditions is recommended not only because shoulder outcome instruments do not adequately capture general health status but also as a tool to compare the outcomes and utility of shoulder diseases and their treatments to other disease processes.

Health utility measures are another option for the measurement of general HRQoL. These measures were developed for use in health economics studies. They judge the patients health status on a scale that includes 1.0 as perfect health and 0.0 as death, however there are conditions that can be negative, as they are considered worse than death from a quality of life standpoint[12,25]. The reason that these are scaled from 0.0 to 1.0 is that this makes calculations for QALY easy for economic analyses and allows comparisons of cost per QALY between different conditions and treatments, for example justifying the cost of rotator cuff repair. Vitale et al[26] showed that the cost per improvement in quality of life from rotator cuff repair was the equivalent of total hip arthroplasty, coronary artery bypass, and more cost efficient than the medical treatment of hypertension. Commonly used health utility measures are the EuroQol 5-domain (EQ-5D) and the Short Form 6D (SF-6D), among others, whose properties are beyond the scope of this review but have been

recently reviewed[15]. Of note, the SF-6D can be calculated if either the SF-36 or the SF-12 is administered as a general HRQoL measure. Health utility measures are an important evaluation tool to include in the functional assessment of shoulder problems if the plan is to understand the financial implications of treatment.

SHOULDER HRQOL MEASURES

While general HRQoL measures are an important part of evaluating patients with shoulder disease, they are not responsive enough to evaluate a patient's overall level of dysfunction in isolation[8]. Some patients with improvement in shoulder function show a decrease in their SF-36 scales after treatment, likely due to the deterioration of other conditions affecting their general health at the same time as their improvement in shoulder pain/function[27]. Because of this, tools evaluating the shoulder or specific disease affecting the shoulder need to be used to complement general HRQoL measures. There have been over thirty shoulder questionnaires described in the literature for evaluation of shoulder pathology[5]. These can be broken down into general shoulder measures and condition (disease) specific shoulder measures.

General shoulder measures are recommended for practice-based evaluation of a heterogeneous group of patients undergoing treatment for shoulder conditions. Condition or disease specific shoulder measures are designed to evaluate homogenous groups of patients with a specific diagnosis and are highly recommended for controlled trials evaluating a specific shoulder disorder. In general, condition specific shoulder measures are less commonly utilized in comparison to the general shoulder measures. The shoulder disorder requiring a condition specific measure the greatest is shoulder instability since many patients with symptomatic shoulder instability have a ceiling effect with general shoulder measures[28]. This section will outline the most commonly utilized general shoulder measures (Table 1) as well as condition specific measures for glenohumeral instability and rotator cuff disease (Table 2).

GENERAL SHOULDER MEASURES

The Constant-Murley score

The Constant-Murley score (CMS) was developed in 1986 and published in 1987 to better estimate the overall functional state of normal and diseased shoulders[29]. A higher score indicates better shoulder function. The CMS continues to be the most commonly reported outcome scale in Europe[11]. The scale combines two fundamentally different metrics: physical examination findings of motion and strength (65 points), and patient-reported subjective evaluation of shoulder function (35 points). In the original description of the CMS, there was no rationale reported for the development and selection of items, or the relative weighting of each component: 15% pain, 20% patient-reported function with activities of daily living, 40% range of motion, and 25% strength testing.

Combining performance-based measures with patient-reported outcomes could be considered an advantage of the CMS; however, it is likely that the reliability of the Constant-Murley score is reduced because patient assessment does not necessarily correlate with objective measurements of shoulder function[30-32]. Still, several studies evaluating the surgical treatment of rotator cuff tears and proximal humerus fractures have found satisfactory correlation between the Constant score and other patient-reported measures[33,34].

The reliability of the Constant score has been questioned with a reported variation between observers as high as 10 units (out of a possible 100)[35]. Conboy *et al*[36] found a low interobserver reliability with 3 different observers evaluating 25 patients using the CMS. On average, these observers differed significantly with regards to total score; the 95% confidence interval that a single measurement represented the true score was 17.7 points. These large, unsatisfactory standard errors contrast the high reliability found in the original publication, where only a 3% interobserver error was reported between 3 observers in 100 abnormal shoulders[29]. Measurement error is most likely attributable to wide variations in strength testing methodology, which was inadequately explained in the original description. Constant *et al*[37] published modifications and guidelines for use of the CMS in 2008 to address these concerns.

Potential advantages of the CMS include its widespread use and prolonged existence, allowing for comparisons across procedure and time. Accordingly, population normative values of the CMS have been established, which aid in score interpretation[38]. Recently, minimum clinically important differences (MCIDs) for the CMS have been reported improving the ability to interpret the clinical relevance of the score as well as design studies using the CMS as the primary outcome tool[39]. The heavy weighting on range of motion and strength may be of benefit when assessing rotator cuff repairs and shoulder arthritis, but has been demonstrated to have problematic ceiling affects in instability patients[36,40]. Reliability, validity, and responsiveness of the CMS are detailed in Table 1.

The University of California Los Angeles shoulder score

The University of California Los Angeles (UCLA) shoulder score was developed in 1981 before modern psychometric development was routinely used[41]. Consequently, the methods utilized in its development are not explained, including question development and weighting. The score is a combination of physical exam findings (active forward elevation and strength) and subjective patient-reported measures (pain, satisfaction, and function). Pain and function are preferentially weighted (20 out of 35 possible points). A higher score indicates better function. The UCLA score has been used to assess a variety of shoulder conditions including total shoulder arthroplasty, rotator cuff repair, and subacromial decompression[42,43].

Limitations of the UCLA are inherent in its design. Many of the questions are double-barreled, meaning that multiple inquiries are combined within a single question.

Table 1 General shoulder measures

Measure	Description	Validity	Reliability	Responsiveness	MCID
The constant score[36,39,74,75]	10 items: Physical Examination (4 motion, 1 strength) Subjective evaluation (1 pain, 4 ADL) Score: 0-100 (Higher = better) 65 points for physical examination 35 points for subjective evaluation	Criterion validity with WORC, Penn, SST, Oxford, and others. Weaker correlation with DASH, ASES, SF-36 Content validity - concern over methods for strength testing Construct validity high except for shoulder instability; scores and strength decrease with age for both sexes	Very good ICC for shoulder dysfunction 0.8-0.87 SEM 8.9	Excellent except for Shoulder instability Effect size: Arthroplasty: 2.23- 3.02 Rotator cuff repair: 1.92 Shoulder instability: 0.20	10.4
UCLA shoulder score[27,65,76-78]	5 items Likert pain scale (1) Function (1) Active forward elevation (1) Forward elevation strength (1) Patient satisfaction (yes/no) Score 0-35 10 pts for pain/function, 5 pts each for active foreward elevation, strength, and satisfaction Can be converted to 0-100 pts for comparison	Criterion validity: Correlated sternly with Constant, ASES, and SF-36; fair to good correlation with SST; fair correlation with constant score; very good correlation with WOSI Construct validity: Demonstrated improvement after subacromial decompression; UCLA score had poor and fair correlation to forward motion and the abduction ratio respectively	Not evaluated	Limited Evaluation Effect size: Subacromial decompression 2.73 at 6 mo Proximal humerus fractures- moderate responsiveness	Not established
DASH[75,79]	30 items Physical activities in arm, shoulder, hand (21) symptoms of pain, tingling, weakness (5) Impact on social activities (4) Score: 0-100 (Lower = Better) Must answer 27 questions to be scored 4 optional sport/music/work items	Criterion validity: Correlated with other scores over different regions of the upper extremity and general outcome measures including the SF-36 Construct validity Difference between: working/not able to work; disease and health state; ability to do what they want versus not able	Excellent ICC: 0.77-0.98 SEM: 2.8-5.2	Excellent Effect size (all studies): 0.4-1.4	10 for shoulder complaints 17 for elbow, wrist and hand
SST[49,68,75]	12 yes/no items	Criterion validity: Strong correlation with ASES, moderately correlated with physical function portion of SF-12 Content validity Differences between: Age groups; shoulder instability versus rotator cuff injury; workers compensation status	Excellent ICC: 0.97-0.99 SEM: N/E	Limited Evaluation Effect size 0.8 in shoulder instability and rotator cuff injuries	2 for rotator cuff disease
ASES evaluation form[55,56,75,80]	11 items Pain VAS (1) Function (10) Score: 0-100 (Higher is better) 50 pts pain/50 pts function Physician assessment is not scored	Criterion validity: Strong correlation with constant-Murley, UCLA, and SST; strong correlation with multiple rotator cuff specific scores; and highly correlated with the SF-12 functional domains, but not the emotional, mental health, and social portions. Content validity Differences found between: Gotten much better and slightly better; minimally, moderately, and maximally functionally limited	Excellent ICC: 0.84-0.96 SEM: 6.7	Excellent Effect size (all studies) 0.9-3.5	6.4 for various shoulder pathologies 12-17 for rotator cuff disease
PENN shoulder score[56,58,81-83]	24 items Pain VAS scales with rest, ADLs, strenuous activities (3) Patient satisfaction VAS (1) Functional assessment section (20) Score 0-100 (Higher = Better) Pain 30 pts Satisfaction 10 pts Function 60 pts	Criterion validity: Excellent correlation with constant; excellent to very good correlation with ASES; Content validity: PSS is negatively affected by chest related, but not other medical comorbidities; pain subscale was not responsive to surgical and nonsurgical treatments	Excellent ICC: 0.94 SEM: 8.5	Not rigorously evaluated Effect size of pain subscale 1.84 for all comers	11.4 for patients with shoulder problems undergoing physical therapy 21 for patients with impingement

UCLA: University of California Los Angeles; ASES: American shoulder and elbow surgeons; DASH: Disabilities of the arm, shoulder and hand; SST: Simple shoulder test; PENN: Pennsylvania; MCID: Minimum clinically important difference: SF-36: Short Form 36; SEM: Standard error of measurement; ICC: Intraclass correlation coefficient; VAS: Visual analog scale; ADL: Activity of daily living.

Table 2 Condition specific shoulder measures

Instability	Description	Validity	Reliability	Responsiveness	MCID
WOSI[62,65,84]	21 items: Physical symptoms (10) Sport/recreation/work function (4) Lifestyle function (4) Emotional function (3) Score: 0-2100 (Lower = Better) (can be converted into 0%-100% scale)	Content validity: Items established by experts and patients Criterion validity: Excellent Correlate: VAS Function and DASH, good with CMS and Rowe	Excellent ICC: 0.87-0.98	Excellent Effect size: 1.67 for stabilization	220/2100
OSIS[28,62]	12 Items: Score: 12-60 (Lower = Better)	Criterion validity: Correlated with rowe and constant scores	Excellent PCC: 0.97	Very good Effect size: 0.8	Not reported
MIIS[62,66]	22 items: Pain (4) Instability (5) Function (8) Occupation and sports (5) Score: 0-100 (lower = better)	Criterion validity: Low to moderate correlation with shoulder rating questionnaire. Otherwise untested	Excellent ICC: 0.98	Not reported	Not reported
Rowe score[63,64]	3 items: Stability (50 points) Motion (20 points) Function (30 points) Score: 0-100 (both subjective and examination dependant)	Content Validity: poorly described development and methodology Criterion Validity: Correlated with WOSI and CMS	Fair ICC 0.7	Very good Effect size 1.2	Not reported
Rotator cuff					
WORC[69]	21 items: Physical symptoms (10 items) Sport/recreation/work function (4 items) Lifestyle function (4 items) Emotional function (3 items) Score: 0-2100 (Lower = Better) (can be converted into 0%-100% scale)	Content validity: Items established by experts and patients Criterion validity: Correlated with ASES, DASH and UCLA	Excellent ICC: 0.96	Excellent Effect size: 0.96	245/2100
RCQoL[85]	34 items: Symptoms and physical complaints (16 items) Sport/recreation (4 items) Work related concerns (4 items) Lifestyle issues (5 items) -Social and Emotional Issues (5 items) Score: 0-3400 (Lower = Worse) (can be converted into 0-100 scale)	Content validity: Items established by experts and patients Criterion validity: Correlated with ASES. Construct validity: able to differentiate large and massive tears	Poor ICC: Not reported Reported as average difference of final score = 5%	Excellent Effect size: Not reported SRM: 1.43	Not reported

WOSI: The Western Ontario Shoulder Instability Index; OSIS: The Oxford Shoulder Instability Score; MIIS: The Melbourne Instability Shoulder Scale; WORC: The Western Ontario rotator cuff index; RCQoL: Rotator cuff quality-of-life measure; MCID: Minimum clinically important difference; ICC: Intra-class correlation coefficient; VAS: Visual analog scale; SRM: Standardized response mean.

For example, the pain scale responses address both frequency of pain along with analgesia type. Respondents might have difficulty picking an appropriate response to the question when they endorse only a portion of one selection, but not the entire response. Furthermore, the satisfaction portion of the instrument only allows for the UCLA score to be logically used post-intervention, making responsiveness impossible to determine. Like the Constant score, including both physical exam and patient self-assessment makes the UCLA multi-dimensional, meaning that it combines multiple domains into a single score. The reliability, validity, and responsiveness are poorly established compared to other outcome measures (Table 1).

Disabilities of the arm, shoulder and hand

The disabilities of the arm, shoulder and hand (DASH) was constructed in 1996 *via* a collaborative effort by the Council of musculoskeletal specialty societies, the American Academy of Orthopaedic Surgeons (AAOS), and the Institute for Work and Health[44]. Sophisticated psychometric techniques were used for item generation to help establish face validity. Lower scores are associated with improved function. The 30-question scale assesses multiple domains including physical function, symptoms, and social/psychological function.

The DASH is intended to measure shoulder, elbow, wrist, and hand function in one combined metric. By design, it does not discriminate between the affected and non-affected extremity. These two properties make the scale more generalizable, but could also be considered an inherent weakness. For example, functional items may not reflect a response to treatment if they mostly involve the dominant arm, especially when the non-dominant are was treated. Despite the possible limitation of a more generalized score, Beaton *et al*[45] found good correlation and responsiveness comparing the DASH and joint-specific measures in a combined population of shoulder, wrist, and hand patients; however, some studies have found only fair responsiveness with the DASH, especially regarding hand conditions[46].

The DASH has been widely studied and offers several advantages. It has been validated in over 15 languages, and normative data has been established for American and Norwegian populations[5,47]. These normative values were 5 for both males and females between the ages of 20 and 29. They increased to 22 and 13 in females and males aged 70 to 79, respectively[47]. The MCID has been reported for both shoulder (MCID = 10) as well as elbow, wrist, and hand patients (MCID = 17)[48]. It is also freely distributed through the AAOS website and has been shown to be valid and reliable for many upper extremity conditions (Table 1). Even though the DASH has been rigorously correlated with shoulder-specific measures and has been shown to have sound psychometric properties, it has not been reported frequently in many shoulder-focused studies. Despite its psychometric properties, shoulder surgeons tend to favor more familiar scales such as the CMS, the American shoulder and Elbow surgeons (ASES), and the simple shoulder test (SST) allowing comparisons of outcomes with prior studies.

The SST

SST was developed in 1992 to reduce responder burden and simplify the process of acquiring outcome information. Questions were developed from: (1) neer's evaluation; (2) ASES evaluation; and (3) Patient complaints and inputs. All twelve questions require yes/no responses. Although this basic format simplifies the survey, the limited range of total points could limit the potential of the SST to detect small but clinically significant changes. The MCID for the SST has been found to be 2 points[49].

The SST has overall sound psychometric properties. Known-group validity tests have shown that the SST can detect differences expected to be observed across different age groups, associated with different shoulder pathologies including instability and rotator cuff tears, and between worker's compensation patients and non-worker's compensation patients[50]. The test is responsive; patients with healed rotator cuff repairs score similarly to normal healthy controls with proven intact rotator cuff tendons by ultrasound[51]. The SST has also been able to distinguish between healthy patients and those with shoulder conditions including osteoarthritis, rheumatoid arthritis, rotator cuff tears, adhesive capsulitis, and instability. The

validity, reliability, and responsiveness of the SST are not as well developed as other measures, but it appears to be psychometrically sound based on available data (Table 1).

The ASES

The ASES score was created by the Society of the American Shoulder and Elbow Surgeons to facilitate standardization of outcome measures and to promote multicenter trials in shoulder and elbow surgery[52]. The ASES score contains a physician-rated and patient-rated section; however, only the pain visual analog scale (VAS) and 10 functional questions are typically used to tabulate the reported ASES score. The total score - 100 maximum points - is weighted 50% for pain and 50% for function. Calculation of the ASES score is somewhat more arduous that other shoulder outcome measures[53]. The final pain score (maximum 50 points) is calculated by subtracting the VAS from 10 and multiplying by five. For the functional portion, each of 10 separate questions is scored on an ordinal scale from 0-3 for a maximal raw functional score of 30 points. The raw score is multiplied by 5/3 to make the maximal functional score out of 50 possible points. The pain and functional portions are then summed to obtain the final ASES score.

Psychometric properties of the ASES have been well established. The validity, reliability, and responsiveness have been assessed in a variety of shoulder problems including: rotator cuff disease, glenohumeral arthritis, shoulder instability, and shoulder arthroplasty[54,55]. The ASES score has also been shown to be valid, reliable, and responsive to non-operative treatments[56]. Minimal clinically important difference for the ASES ranges from 6.4 for various shoulder disorders to 12-17 points - depending on confidence level - in rotator cuff problems[49,56]. The ASES score has been translated into German and validated, but is not available in as many languages as the DASH. Correlation with other shoulder and upper extremity measures is high for the ASES score (Table 1).

Although the ASES score has been rigorously evaluated, some inherent limitations are noteworthy. Weighting of the ASES score favors the domains of pain and patient-reported function. Unlike the Constant-Murley score, physician assessment is not included in the final score. This could be considered both a strength and weakness of the ASES, but it should be noted in interpreting results. The shoulder instability VAS of the ASES has been removed in some versions, although the scale has still been responsive to instability treatments without this portion of the survey[55]. A final limitation is that higher functioning patients may experience ceiling effects due to the response structure[57].

Pennsylvania shoulder score

The pennsylvania shoulder score (PSS) is a 100-point shoulder specific scale comprised of pain (30%), satisfaction (10%), and function (60%). There are three pain VAS scores: one each for pain at rest, pain with everyday activities, and pain with strenuous activities. Patient satisfaction is determined from 0-10 on numeric rating scale.

The remaining functional portion of the scale is comprised of 20 questions with maximal ordinal responses that are assigned a maximal value of 3 points each.

The psychometric properties of the PSS appear favorable, but this scale has not been rigorously tested in multiple investigations. Leggin *et al*[58] performed the most thorough assessment of the PSS finding overall good reliability, internal consistency, and correlation with the ASES and CMS. The PSS was found to be responsive, and had an MCID of 11.4 in patients that underwent non-operative treatment of various shoulder problems.

The PSS has been used less frequently in the literature compared to other outcome scales discussed in this article. It seems that this outcome measure has been embraced regionally in the United States, with the majority of studies using the Penn shoulder scale originating out of only a handful of institutions. The PSS was used in one study that recommended against augmentation of large and massive rotator cuff tears with porcine xenografts[59]. Proximal humerus fractures, latissimus tendon transfers for irreparable rotator cuff, and non-operative therapies have all been evaluated with the PSS[58,60,61].

CONDITION SPECIFIC SHOULDER MEASURES

Instability
Instability is the most common diagnosis in which condition specific measures are used. The presentation of patients with symptomatic instability is different from other shoulder pathology. After reduction of the acute dislocation, patients with symptomatic instability requiring treatment commonly present with recurrent instability or apprehension, not pain and decreased function as is more common with other shoulder diagnoses. This leads to poor responsiveness and significant ceiling effects when general shoulder measures are used for patients with instability[28]. Because of this, specific instability scores have been developed to study shoulder instability that are more responsive to treatment effects[62]. The most common validated patient reported outcome measures for shoulder instability are the Western Ontario Shoulder Instability Index (WOSI), the Oxford Shoulder Instability Score (OSIS), and the Melbourne Instability Shoulder Scale (MISS). However, the most commonly used evaluation is the Rowe score, which was also the first shoulder score described in 1978. The Rowe score, similar to the UCLA shoulder score, was first described before modern psychometric development was implemented limiting it's psychometric properties[63,64]. The WOSI, MISS and OSIS have been developed with recent psychometric evaluations[28,65,66]. The properties of these scores are described in Table 2. The WOSI is more responsive to treatment of instability than the Rowe score in patients both non-operatively and operatively treated for traumatic instability[65,67]. Overall, the WOSI has the strongest psychometric properties and has undergone the most rigorous testing despite the fact that the Rowe is the most commonly re-

ported instability measure. Based upon the strength of its psychometric properties, the WOSI is the recommended condition specific instrument for shoulder instability.

Rotator cuff
There had also been evaluation tools designed specifically for the evaluation of patients with rotator cuff disease. The two most common rotator cuff specific tools are the Western ontario rotator cuff Index (WORC) and the rotator cuff quality-of-life measure (RCQoL). General shoulder measures are commonly used for patients with rotator cuff disease as well and these have been shown to be valid and responsive in this patient population[55,68]. Because of the utility of other general shoulder instruments the need for specific rotator cuff instruments is called into question. Overall, generalized shoulder instruments do not show the same kind of ceiling effect with rotator cuff disease that they do with instability. Again, the WORC has the strongest psychometric properties and has undergone the most rigorous testing[69]. This makes it the instrument of choice if a condition specific measure for rotator cuff disease is desired. The properties of these two scores are presented in Table 2.

COMPUTER ADAPTIVE TESTING
The National Institutes of Health Roadmap Initiative has recently launched the Patient-Reported Outcomes Measurement Information System (PROMIS) that is available for clinical use for a variety of health domains, including physical function[70]. This novel instrument was developed to: (1) obtain precise estimations of specific health-related domains; (2) eliminate floor and ceiling effects by validating a large "bank" of questions; and (3) reduce patient respondent burden by minimizing the number of questions (typically only 3-5)[71]. PROMIS is made possible using computerized adaptive testing (CAT), which takes each individual's previous answer into account when asking subsequent questions. By asking "intelligent" questions - *i.e.*, it is unnecessary to ask if a patient can comb their hair if they can throw a baseball - precise results can be achieved with only a few questions selected from a large item bank[72]. Therefore, different sets of questions will be administered to different individuals with the results reported on a common scale. This approach differs from classical test theory, where all (or nearly all) questions included in the static survey must be answered to use the metric[44].

The PROMIS Physical Function CAT (PF-CAT) is designed to measure a single domain. This contrasts with commonly used shoulder scales such as the ASES, CMS, DASH, and UCLA that lump multiple domains (pain, physical function, and objective tests) into a single scale. This can be considered both an advantage and disadvantage; however, if desired, CAT tests that measure pain, anxiety, and depression are also available for administration. One concern regarding the PF-CAT is that it includes questions on both the upper and lower extremi-

ties that could limit the responsiveness of the metric. To address this concern, an upper extremity CAT (UE CAT) has been developed and has been shown to correlate strongly with the DASH in non-shoulder upper extremity patients[73]. An upper extremity specific CAT could eliminate small ceiling effects that were found when assessing the PF CAT in some upper extremity patients[72].

In general, the psychometric properties of the PF and UE CAT have not yet been rigorously evaluated. The potential benefits of CAT testing include: reduced time to completion and decreased patient responder-burden; reducing or eliminating floor and ceiling effects; unidimensionality that could clarify interpretation of results; and the ability to add or subtract questions from the item bank without the need to recreate and validate an entirely new scale. It is likely that PROMIS PROs will be reported in studies evaluating shoulder outcomes going forward, and therefore the reader should become aware of this methodology.

CONCLUSION

A variety of outcome assessment tools can be utilized to evaluate patients with shoulder disorders including general HRQoL measures, health utility measures, general shoulder HRQoL measures and, in the setting of instability, condition specific shoulder measures. The SF-36 and SF-12 are the most validated and commonly used general HRQoL measures in the orthopaedic literature. Utilizing one of these also allows for calculation of the SF-5D as a health utility measure for economic analysis. There are multiple general shoulder measures that are acceptable for use as a general shoulder measure including the ASES score, SST and CMS, however in the setting of instability it is recommended to use a condition specific measure due to the ceiling effects of general shoulder measures. The WOSI is the most rigorously tested and validated of the instability measures. Finally, computer adaptive testing and the PROMIS database is emerging as a unique and powerful tool in evaluating both general and joint specific HRQoL that may allow for more efficient evaluation of patient outcomes in the near future.

REFERENCES

1   **Bayley KB**, London MR, Grunkemeier GL, Lansky DJ. Measuring the success of treatment in patient terms. *Med Care* 1995; **33**: AS226-AS235 [PMID: 7723451]

2   **Brokelman RB**, van Loon CJ, Rijnberg WJ. Patient versus surgeon satisfaction after total hip arthroplasty. *J Bone Joint Surg Br* 2003; **85**: 495-498 [PMID: 12793551]

3   **Noble PC**, Fuller-Lafreniere S, Meftah M, Dwyer MK. Challenges in outcome measurement: discrepancies between patient and provider definitions of success. *Clin Orthop Relat Res* 2013; **471**: 3437-3445 [PMID: 23955192 DOI: 10.1007/s11999-013-3198-x]

4   **Gartland JJ**. Orthopaedic clinical research. Deficiencies in experimental design and determinations of outcome. *J Bone Joint Surg Am* 1988; **70**: 1357-1364 [PMID: 3182887]

5   **Wright RW**, Baumgarten KM. Shoulder outcomes measures. *J Am Acad Orthop Surg* 2010; **18**: 436-444 [PMID: 20595136]

6   **Brophy RH**, Beauvais RL, Jones EC, Cordasco FA, Marx RG. Measurement of shoulder activity level. *Clin Orthop Relat Res* 2005; **439**: 101-108 [PMID: 16205147 DOI: 10.1097/01.blo.0000173255.85016.1f]

7   **Hepper CT**, Smith MV, Steger-May K, Brophy RH. Normative data of shoulder activity level by age and sex. *Am J Sports Med* 2013; **41**: 1146-1151 [PMID: 23524150 DOI: 10.1177/0363546513480613]

8   **Patel AA**, Donegan D, Albert T. The 36-item short form. *J Am Acad Orthop Surg* 2007; **15**: 126-134 [PMID: 17277259]

9   **Rihn JA**, Currier BL, Phillips FM, Glassman SD, Albert TJ. Defining the value of spine care. *J Am Acad Orthop Surg* 2013; **21**: 419-426 [PMID: 23818029 DOI: 10.5435/JAAOS-21-07-419]

10   **Smith MV**, Calfee RP, Baumgarten KM, Brophy RH, Wright RW. Upper extremity-specific measures of disability and outcomes in orthopaedic surgery. *J Bone Joint Surg Am* 2012; **94**: 277-285 [PMID: 22298061 DOI: 10.2106/JBJS.J.01744]

11   **Kirkley A**, Griffin S, Dainty K. Scoring systems for the functional assessment of the shoulder. *Arthroscopy* 2003; **19**: 1109-1120 [PMID: 14673454 DOI: 10.1016/j.arthro.2003.10.030]

12   **Brinker MR**, O'Connor DP. Stakeholders in outcome measures: review from a clinical perspective. *Clin Orthop Relat Res* 2013; **471**: 3426-3436 [PMID: 24072624 DOI: 10.1007/s11999-013-3265-3]

13   **McHorney CA**, Ware JE, 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**: 247-263 [PMID: 8450681 DOI: 10.1097/00005650-199303000-00006]

14   **Ware JE**, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; **30**: 473-483 [PMID: 1593914 DOI: 10.1097/00005650-199206000-00002]

15   **Busija L**, Pausenberger E, Haines TP, Haymes S, Buchbinder R, Osborne RH. Adult measures of general health and health-related quality of life: Medical Outcomes Study Short Form 36-Item (SF-36) and Short Form 12-Item (SF-12) Health Surveys, Nottingham Health Profile (NHP), Sickness Impact Profile (SIP), Medical Outcomes Study Short Form 6D (SF-6D), Health Utilities Index Mark 3 (HUI3), Quality of Well-Being Scale (QWB), and Assessment of Quality of Life (AQoL). *Arthritis Care Res* (Hoboken) 2011; **63** Suppl 11: S383-S412 [PMID: 22588759 DOI: 10.1002/acr.20541]

16   **Ware J**, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996; **34**: 220-233 [PMID: 8628042 DOI: 10.1097/00005650-199603000-00003]

17   **Garratt A**, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. *BMJ* 2002; **324**: 1417 [PMID: 12065262 DOI: 10.1136/bmj.324.7351.1417]

18   **Castricini R**, Gasparini G, Di Luggo F, De Benedetto M, De Gori M, Galasso O. Health-related quality of life and functionality after reverse shoulder arthroplasty. *J Shoulder Elbow Surg* 2013; **22**: 1639-1649 [PMID: 23523306 DOI: 10.1016/j.jse.2013.01.020]

19   **Yoo JH**, Cho NS, Rhee YG. Effect of postoperative repair integrity on health-related quality of life after rotator cuff repair: healed versus retear group. *Am J Sports Med* 2013; **41**: 2637-2644 [PMID: 23942286 DOI: 10.1177/0363546513499152]

20   **Carter MJ**, Mikuls TR, Nayak S, Fehringer EV, Michaud K. Impact of total shoulder arthroplasty on generic and shoulder-specific health-related quality-of-life measures: a systematic literature review and meta-analysis. *J Bone Joint Surg Am* 2012; **94**: e127 [PMID: 22992856 DOI: 10.2106/JBJS.K.00204]

21   **Gartsman GM**, Brinker MR, Khan M, Karahan M. Self-assessment of general health status in patients with five common shoulder conditions. *J Shoulder Elbow Surg* 1998; **7**: 228-237 [PMID: 9658347 DOI: 10.1016/S1058-2746(98)90050-7]

22   **Boorman RS**, Kopjar B, Fehringer E, Churchill RS, Smith

K, Matsen FA. The effect of total shoulder arthroplasty on self-assessed health status is comparable to that of total hip arthroplasty and coronary artery bypass grafting. *J Shoulder Elbow Surg* 2003; **12**: 158-163 [PMID: 12700569 DOI: 10.1067/mse.2003.18]

23 **Beaton DE**, Richards RR. Measuring function of the shoulder. A cross-sectional comparison of five questionnaires. *J Bone Joint Surg Am* 1996; **78**: 882-890 [PMID: 8666606]

24 **Oh JH**, Jo KH, Kim WS, Gong HS, Han SG, Kim YH. Comparative evaluation of the measurement properties of various shoulder outcome instruments. *Am J Sports Med* 2009; **37**: 1161-1168 [PMID: 19403837 DOI: 10.1177/0363546508330135]

25 **Rudmik L**, Drummond M. Health economic evaluation: important principles and methodology. *Laryngoscope* 2013; **123**: 1341-1347 [PMID: 23483522 DOI: 10.1002/lary.23943]

26 **Vitale MA**, Vitale MG, Zivin JG, Braman JP, Bigliani LU, Flatow EL. Rotator cuff repair: an analysis of utility scores and cost-effectiveness. *J Shoulder Elbow Surg* 2007; **16**: 181-187 [PMID: 17399623 DOI: 10.1016/j.jse.2006.06.013]

27 **Gartsman GM**, Brinker MR, Khan M. Early effectiveness of arthroscopic repair for full-thickness tears of the rotator cuff: an outcome analysis. *J Bone Joint Surg Am* 1998; **80**: 33-40 [PMID: 9469306]

28 **Dawson J**, Fitzpatrick R, Carr A. The assessment of shoulder instability. The development and validation of a questionnaire. *J Bone Joint Surg Br* 1999; **81**: 420-426 [PMID: 10872358 DOI: 10.1302/0301-620X.81B3.9044]

29 **Constant CR**, Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop Relat Res* 1987; **(214)**: 160-164 [PMID: 3791738]

30 **Namdari S**, Donegan RP, Chamberlain AM, Galatz LM, Yamaguchi K, Keener JD. Factors affecting outcome after structural failure of repaired rotator cuff tears. *J Bone Joint Surg Am* 2014; **96**: 99-105 [PMID: 24430408 DOI: 10.2106/JBJS.M.00551]

31 **Galatz LM**, Ball CM, Teefey SA, Middleton WD, Yamaguchi K. The outcome and repair integrity of completely arthroscopically repaired large and massive rotator cuff tears. *J Bone Joint Surg Am* 2004; **86-A**: 219-224 [PMID: 14960664]

32 **Boileau P**, Brassart N, Watkinson DJ, Carles M, Hatzidakis AM, Krishnan SG. Arthroscopic repair of full-thickness tears of the supraspinatus: does the tendon really heal? *J Bone Joint Surg Am* 2005; **87**: 1229-1240 [PMID: 15930531 DOI: 10.2106/JBJS.D.02035]

33 **Allom R**, Colegate-Stone T, Gee M, Ismail M, Sinha J. Outcome analysis of surgery for disorders of the rotator cuff: a comparison of subjective and objective scoring tools. *J Bone Joint Surg Br* 2009; **91**: 367-373 [PMID: 19258614 DOI: 10.1302/0301-620X.91B3.20714]

34 **Baker P**, Nanda R, Goodchild L, Finn P, Rangan A. A comparison of the Constant and Oxford shoulder scores in patients with conservatively treated proximal humeral fractures. *J Shoulder Elbow Surg* 2008; **17**: 37-41 [PMID: 18036850 DOI: 10.1016/j.jse.2007.04.019]

35 **Rocourt MH**, Radlinger L, Kalberer F, Sanavi S, Schmid NS, Leunig M, Hertel R. Evaluation of intratester and intertester reliability of the Constant-Murley shoulder assessment. *J Shoulder Elbow Surg* 2008; **17**: 364-369 [PMID: 18329560 DOI: 10.1016/j.jse.2007.06.024]

36 **Conboy VB**, Morris RW, Kiss J, Carr AJ. An evaluation of the Constant-Murley shoulder assessment. *J Bone Joint Surg Br* 1996; **78**: 229-232 [PMID: 8666631]

37 **Constant CR**, Gerber C, Emery RJ, Søjbjerg JO, Gohlke F, Boileau P. A review of the Constant score: modifications and guidelines for its use. *J Shoulder Elbow Surg* 2008; **17**: 355-361 [PMID: 18218327 DOI: 10.1016/j.jse.2007.06.022]

38 **Katolik LI**, Romeo AA, Cole BJ, Verma NN, Hayden JK, Bach BR. Normalization of the Constant score. *J Shoulder Elbow Surg* 2005; **14**: 279-285 [PMID: 15889027 DOI: 10.1016/j.jse.2004.10.009]

39 **Kukkonen J**, Kauko T, Vahlberg T, Joukainen A, Äärimaa V. Investigating minimal clinically important difference for Constant score in patients undergoing rotator cuff surgery. *J Shoulder Elbow Surg* 2013; **22**: 1650-1655 [PMID: 23850308 DOI: 10.1016/j.jse.2013.05.002]

40 **Kemp KA**, Sheps DM, Beaupre LA, Styles-Tripp F, Luciak-Corea C, Balyk R. An evaluation of the responsiveness and discriminant validity of shoulder questionnaires among patients receiving surgical correction of shoulder instability. *ScientificWorldJournal* 2012; **2012**: 410125 [PMID: 23002386 DOI: 10.1100/2012/410125]

41 **Amstutz HC**, Sew Hoy AL, Clarke IC. UCLA anatomic total shoulder arthroplasty. *Clin Orthop Relat Res* 1981; **(155)**: 7-20 [PMID: 7226634]

42 **Fealy S**, Kingham TP, Altchek DW. Mini-open rotator cuff repair using a two-row fixation technique: outcomes analysis in patients with small, moderate, and large rotator cuff tears. *Arthroscopy* 2002; **18**: 665-670 [PMID: 12098132 DOI: 10.1053/jars.2002.32589]

43 **O'Connor DA**, Chipchase LS, Tomlinson J, Krishnan J. Arthroscopic subacromial decompression: responsiveness of disease-specific and health-related quality of life outcome measures. *Arthroscopy* 1999; **15**: 836-840 [PMID: 10564861 DOI: 10.1053/ar.1999.v15.015083]

44 **Hudak PL**, Amadio PC, Bombardier C. 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**: 602-608 [PMID: 8773720 DOI: 10.1002/(SICI)1097-0274(199606)29:]

45 **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**: 128-146 [PMID: 11382253 DOI: 10.1016/S0894-1130(01)80043-0]

46 **Gabel CP**, Michener LA, Burkett B, Neller A. The Upper Limb Functional Index: development and determination of reliability, validity, and responsiveness. *J Hand Ther* 2006; **19**: 328-348; quiz 349 [PMID: 16861132 DOI: 10.1197/j.jht.2006.04.001]

47 **Aasheim T**, Finsen V. The DASH and the QuickDASH instruments. Normative values in the general population in Norway. *J Hand Surg Eur Vol* 2014; **39**: 140-144 [PMID: 23520389 DOI: 10.1177/1753193413481302]

48 **Schmitt JS**, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *J Clin Epidemiol* 2004; **57**: 1008-1018 [PMID: 15528051 DOI: 10.1016/j.jclinepi.2004.02.007]

49 **Tashjian RZ**, Deloach J, Green A, Porucznik CA, Powell AP. Minimal clinically important differences in ASES and simple shoulder test scores after nonoperative treatment of rotator cuff disease. *J Bone Joint Surg Am* 2010; **92**: 296-303 [PMID: 20124055 DOI: 10.2106/JBJS.H.01296]

50 **Henn RF**, Tashjian RZ, Kang L, Green A. Patients with workers' compensation claims have worse outcomes after rotator cuff repair. *J Bone Joint Surg Am* 2008; **90**: 2105-2113 [PMID: 18829907 DOI: 10.2106/JBJS.F.00260]

51 **Fehringer EV**, Sun J, Cotton J, Carlson MJ, Burns EM. Healed cuff repairs impart normal shoulder scores in those 65 years of age and older. *Clin Orthop Relat Res* 2010; **468**: 1521-1525 [PMID: 19763715 DOI: 10.1007/s11999-009-1103-4]

52 **Richards RR**, An KN, Bigliani LU, Friedman RJ, Gartsman GM, Gristina AG, Iannotti JP, Mow VC, Sidles JA, Zuckerman JD. A standardized method for the assessment of shoulder function. *J Shoulder Elbow Surg* 1994; **3**: 347-352 [PMID: 22958838 DOI: 10.1016/S1058-2746(09)80019-0]

53 **Bot SD**, Terwee CB, van der Windt DA, Bouter LM, Dekker J, de Vet HC. Clinimetric evaluation of shoulder disability

questionnaires: a systematic review of the literature. *Ann Rheum Dis* 2004; **63**: 335-341 [PMID: 15020324 DOI: 10.1136/ard.2003.007724]

54   **Angst F**, Pap G, Mannion AF, Herren DB, Aeschlimann A, Schwyzer HK, Simmen BR. Comprehensive assessment of clinical outcome and quality of life after total shoulder arthroplasty: usefulness and validity of subjective outcome measures. *Arthritis Rheum* 2004; **51**: 819-828 [PMID: 15478159 DOI: 10.1002/art.20688]

55   **Kocher MS**, Horan MP, Briggs KK, Richardson TR, O'Holleran J, Hawkins RJ. Reliability, validity, and responsiveness of the American Shoulder and Elbow Surgeons subjective shoulder scale in patients with shoulder instability, rotator cuff disease, and glenohumeral arthritis. *J Bone Joint Surg Am* 2005; **87**: 2006-2011 [PMID: 16140816 DOI: 10.2106/JBJS.C.01624]

56   **Michener LA**, McClure PW, Sennett BJ. American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form, patient self-report section: reliability, validity, and responsiveness. *J Shoulder Elbow Surg* 2002; **11**: 587-594 [PMID: 12469084 DOI: 10.1067/mse.2002.127096]

57   **Bryant D**, Litchfield R, Sandow M, Gartsman GM, Guyatt G, Kirkley A. A comparison of pain, strength, range of motion, and functional outcomes after hemiarthroplasty and total shoulder arthroplasty in patients with osteoarthritis of the shoulder. A systematic review and meta-analysis. *J Bone Joint Surg Am* 2005; **87**: 1947-1956 [PMID: 16140808 DOI: 10.2106/JBJS.D.02854]

58   **Leggin BG**, Michener LA, Shaffer MA, Brenneman SK, Iannotti JP, Williams GR. The Penn shoulder score: reliability and validity. *J Orthop Sports Phys Ther* 2006; **36**: 138-151 [PMID: 16596890 DOI: 10.2519/jospt.2006.36.3.138]

59   **Iannotti JP**, Codsi MJ, Kwon YW, Derwin K, Ciccone J, Brems JJ. Porcine small intestine submucosa augmentation of surgical repair of chronic two-tendon rotator cuff tears. A randomized, controlled trial. *J Bone Joint Surg Am* 2006; **88**: 1238-1244 [PMID: 16757756 DOI: 10.2106/JBJS.E.00524]

60   **Ricchetti ET**, Warrender WJ, Abboud JA. Use of locking plates in the treatment of proximal humerus fractures. *J Shoulder Elbow Surg* 2010; **19**: 66-75 [PMID: 20188270 DOI: 10.1016/j.jse.2010.01.001]

61   **Codsi MJ**, Hennigan S, Herzog R, Kella S, Kelley M, Leggin B, Williams GR, Iannotti JP. Latissimus dorsi tendon transfer for irreparable posterosuperior rotator cuff tears. Surgical technique. *J Bone Joint Surg Am* 2007; **89** Suppl 2 Pt.1: 1-9 [PMID: 17332121 DOI: 10.2106/JBJS.F.01160]

62   **Rouleau DM**, Faber K, MacDermid JC. Systematic review of patient-administered shoulder functional scores on instability. *J Shoulder Elbow Surg* 2010; **19**: 1121-1128 [PMID: 21070955 DOI: 10.1016/j.jse.2010.07.003]

63   **Rowe CR**, Patel D, Southmayd WW. The Bankart procedure: a long-term end-result study. *J Bone Joint Surg Am* 1978; **60**: 1-16 [PMID: 624747]

64   **Skare Ø**, Schrøder CP, Mowinckel P, Reikerås O, Brox JI. Reliability, agreement and validity of the 1988 version of the Rowe Score. *J Shoulder Elbow Surg* 2011; **20**: 1041-1049 [PMID: 21831671 DOI: 10.1016/j.jse.2011.04.024]

65   **Kirkley A**, Griffin S, McLintock H, Ng L. The development and evaluation of a disease-specific quality of life measurement tool for shoulder instability. The Western Ontario Shoulder Instability Index (WOSI). *Am J Sports Med* 1998; **26**: 764-772 [PMID: 9850776]

66   **Watson L**, Story I, Dalziel R, Hoy G, Shimmin A, Woods D. A new clinical outcome measure of glenohumeral joint instability: the MISS questionnaire. *J Shoulder Elbow Surg* 1998; **14**: 22-30 [PMID: 15723010 DOI: 10.1016/j.jse.2004.05.002]

67   **Kirkley A**, Griffin S, Richards C, Miniaci A, Mohtadi N. Prospective randomized clinical trial comparing the effectiveness of immediate arthroscopic stabilization versus immobilization and rehabilitation in first traumatic anterior dislocations of the shoulder. *Arthroscopy* 1999; **15**: 507-514 [PMID: 10424554 DOI: 10.1053/ar.1999.v15.015050]

68   **Godfrey J**, Hamman R, Lowenstein S, Briggs K, Kocher M. Reliability, validity, and responsiveness of the simple shoulder test: psychometric properties by age and injury type. *J Shoulder Elbow Surg* 2007; **16**: 260-267 [PMID: 17188906 DOI: 10.1016/j.jse.2006.07.003]

69   **de Witte PB**, Henseler JF, Nagels J, Vliet Vlieland TP, Nelissen RG. The Western Ontario rotator cuff index in rotator cuff disease patients: a comprehensive reliability and responsiveness validation study. *Am J Sports Med* 2012; **40**: 1611-1619 [PMID: 22582227 DOI: 10.1177/0363546512446591]

70   **Cella D**, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, Ader D, Fries JF, Bruce B, Rose M. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care* 2007; **45**: S3-S11 [PMID: 17443116 DOI: 10.1097/01.mlr.0000258615.42478.55]

71   **Chakravarty EF**, Bjorner JB, Fries JF. Improving patient reported outcomes using item response theory and computerized adaptive testing. *J Rheumatol* 2007; **34**: 1426-1431 [PMID: 17552069]

72   **Hung M**, Clegg DO, Greene T, Saltzman CL. Evaluation of the PROMIS physical function item bank in orthopaedic patients. *J Orthop Res* 2011; **29**: 947-953 [PMID: 21437962 DOI: 10.1002/jor.21308]

73   **Döring AC**, Nota SP, Hageman MG, Ring DC. Measurement of upper extremity disability using the Patient-Reported Outcomes Measurement Information System. *J Hand Surg Am* 2014; **39**: 1160-1165 [PMID: 24799143 DOI: 10.1016/j.jhsa.2014.03.013]

74   **Roy JS**, MacDermid JC, Woodhouse LJ. A systematic review of the psychometric properties of the Constant-Murley score. *J Shoulder Elbow Surg* 2010; **19**: 157-164 [PMID: 19559630 DOI: 10.1016/j.jse.2009.04.008]

75   **Roy JS**, MacDermid JC, Woodhouse LJ. Measuring shoulder function: a systematic review of four questionnaires. *Arthritis Rheum* 2009; **61**: 623-632 [PMID: 19405008 DOI: 10.1002/art.24396]

76   **Romeo AA**, Mazzocca A, Hang DW, Shott S, Bach BR. Shoulder scoring scales for the evaluation of rotator cuff repair. *Clin Orthop Relat Res* 2004; **(427)**: 107-114 [PMID: 15552145 DOI: 10.1097/01.blo.0000142624.05526.dd]

77   **Roddey TS**, Olson SL, Cook KF, Gartsman GM, Hanten W. Comparison of the University of California-Los Angeles Shoulder Scale and the Simple Shoulder Test with the shoulder pain and disability index: single-administration reliability and validity. *Phys Ther* 2000; **80**: 759-768 [PMID: 10911414]

78   **van de Water AT**, Shields N, Davidson M, Evans M, Taylor NF. Reliability and validity of shoulder function outcome measures in people with a proximal humeral fracture. *Disabil Rehabil* 2014; **36**: 1072-1079 [PMID: 24001265 DOI: 10.3109/09638288.2013.829529]

79   **Franchignoni F**, Vercelli S, Giordano A, Sartorio F, Bravini E, Ferriero G. Minimal clinically important difference of the disabilities of the arm, shoulder and hand outcome measure (DASH) and its shortened version (QuickDASH). *J Orthop Sports Phys Ther* 2014; **44**: 30-39 [PMID: 24175606 DOI: 10.2519/jospt.2014.4893]

80   **Skutek M**, Fremerey RW, Zeichen J, Bosch U. Outcome analysis following open rotator cuff repair. Early effectiveness validated using four different shoulder assessment scales. *Arch Orthop Trauma Surg* 2000; **120**: 432-436 [PMID: 10968533 DOI: 10.1007/s004020000133]

81   **Wylie JD**, Bershadsky B, Iannotti JP. The effect of medical comorbidity on self-reported shoulder-specific health related quality of life in patients with shoulder disease. *J Shoulder Elbow Surg* 2010; **19**: 823-828 [PMID: 20303289 DOI: 10.1016/j.jse.2009.11.052]

82   **Michener LA**, Snyder AR, Leggin BG. Responsiveness of the numeric pain rating scale in patients with shoulder pain and the effect of surgical status. *J Sport Rehabil* 2011; **20**: 115-128

[PMID: 21411827]

83   **Michener LA**, Snyder Valier AR, McClure PW. Defining substantial clinical benefit for patient-rated outcome tools for shoulder impingement syndrome. *Arch Phys Med Rehabil* 2013; **94**: 725-730 [PMID: 23164979 DOI: 10.1016/j.apmr.2012.11.011]

84   **Angst F**, Schwyzer HK, Aeschlimann A, Simmen BR, Goldhahn J. Measures of adult shoulder function: Disabilities of the Arm, Shoulder, and Hand Questionnaire (DASH) and its short version (QuickDASH), Shoulder Pain and Disability Index (SPADI), American Shoulder and Elbow Surgeons (ASES) Society standardized shoulder assessment form, Constant (Murley) Score (CS), Simple Shoulder Test (SST), Oxford Shoulder Score (OSS), Shoulder Disability Questionnaire (SDQ), and Western Ontario Shoulder Instability Index (WOSI). *Arthritis Care Res (Hoboken)* 2011; **63** Suppl 11: S174-S188 [PMID: 22588743 DOI: 10.1002/acr.20630]

85   **Hollinshead RM**, Mohtadi NG, Vande Guchte RA, Wadey VM. Two 6-year follow-up studies of large and massive rotator cuff tears: comparison of outcome measures. *J Shoulder Elbow Surg* 2000; **9**: 373-381 [PMID: 11075319 DOI: 10.1067/mse.2000.108389]

Baishideng®

Published by **Baishideng Publishing Group Inc**
8226 Regency Drive, Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com