A Systematic Review of Outcome Measures Assessing Disability Following Upper Extremity Trauma

OBJECTIVES: To define upper extremity outcome measures focusing on trauma and level of initial psychometric evaluation and to assess methodological quality of relevant patient-reported outcome (PRO) measures.

DATA SOURCES: A broad search strategy using PubMed, OVID, CINAHL, and PsycINFO was deployed and reported using PRISMA (PROSPERO: CRD42016046243).

STUDY SELECTION, EXTRACTION, SYNTHESIS: PRO measures involving orthopedic trauma in their original development were selected and original publications assessed, including psychometric evaluations. Extraction, synthesis, and quality assessment were performed using COSMIN.

RESULTS: Of 144 upper extremity outcome measures, the majority were designed for the shoulder, wrist, and hand; 20% (n = 29/144) involved trauma conditions in their initial development, PRO measurements, and psychometric evaluation on introduction. Methodological quality was highly variable.

CONCLUSION: A few PRO measures were originally designed for use in upper extremity trauma. Methodological quality and psychometric evaluation need to improve. This review aims to highlight strengths and weaknesses and guide decision making in this field.

Outcome measurement in orthopaedics has evolved rapidly over the past 20 years, and there are many patient-reported or clinician-based outcome measures. The popularity of patient-reported outcome (PRO) measurement, in particular, has grown in response to the perception that clinicians have an incomplete understanding of the true impact of disease on a patient’s life and the complexity of the human illness experience. PRO measures, by definition, focus on quantifying the subjective impact of health from the patient’s perspective, commonly referred to as “disability” in contrast to “impairment” (objective pathophysiology). Common orthopaedic outcomes such as range of motion and fracture union represent the biomedical paradigm. PRO measures represent the biopsychosocial paradigm (including the influence of thoughts, emotions, behaviors, and circumstances) on symptoms and limitation.

The International Classification of Functioning, Disability and Health defines disability as a multidimensional concept related to the dynamic interaction between body functions...
and structures, activity limitations, and participation restrictions alongside environmental and personal factors. These components are influenced by impairment (ie, problems with structure and function of the body leading to significant deviation and loss), psychosocial factors, and symptom experience. The alleviation of disability, in this wider context, is the primary aim of most orthopaedic interventions.

Orthopaedic trauma is often associated with a significant impact on the magnitude of disability and the factors influencing it, which can affect an individual's quality of life in several health domains. There is increasing evidence that disability is less associated with measures of impairment and objective pathophysiology than the subjective psychosocial aspects of illness. Factors likely to mediate these interactions include anxiety, depression, ineffective coping, pain catastrophizing, and kinesiophobia, as well as social status, support, financial loss, and secondary gain. This has an influence on recovery following musculoskeletal trauma, which is shown to have a stronger association with pain intensity and disability than biomedical factors, such as fracture type. Furthermore, studies such as those conducted by Bhandari et al reported on a significant number of patients experiencing orthopaedic trauma breach thresholds for psychological distress.

Upper limb injuries demonstrate reduced health-related quality of life indices compared with trauma involving other regions. The inability to feed, clothe, and care for oneself following injury, particularly involving a dominant arm, can be extremely debilitating. A study involving proximal humerus fractures demonstrated that measures of impairment, such as range of motion...
Table 1
Outcome Measures for Upper Extremity Conditions Classified by Conditions, Instrument, and Initial Psychometric Evaluation

| Region | Outcome Measure | Short Title | Clinical Conditions in Index Evaluation Study | Instrument Classification | Psychometric Evaluation |
|--------|----------------|-------------|-----------------------------------------------|---------------------------|-------------------------|
| Arm    | Disabilities of the Arm, Shoulder, and Hand (1996) | DASH Combo | Colles fracture; humerus fracture; thumb CMCJ arthritis; CTS; rotator cuff tendinopathy; lateral epicondylitis; de Quervain tenosynovitis; OA; RhA; nonspecific; other | Multiregion PRO | P | Staged psychometric evaluation |
| QuickDASH (2005) | QuickDASH Combo | Colles fracture; humerus fracture; thumb CMCJ arthritis; CTS; rotator cuff tendinopathy; lateral epicondylitis; de Quervain tenosynovitis; OA; RhA; nonspecific; other | Multiregion PRO | P | Psychometric evaluation and validation versus DASH; VAS ability to function in daily activities, rating of problem, pain severity, ability to work. |
| Upper Extremity Function Test (1965) | UEFT Combo | Upper extremities with traumatic, neurological and arthritic impairments; amputations | Multiregion PRO | P | Psychometric evaluation and validation versus hand activities of daily living |
| Upper Extremity Functional Index (2001) | UEFI nos | Upper extremity dysfunction nos | Multiregion PRO | P | Psychometric evaluation and validation versus UEFS |
| Upper Extremity Functional Limitation Scale (2001) | UEFLS nos | Elderly women with difficulty performing upper extremity tasks | Multiregion PRO | P | Psychometric evaluation and validation versus fingers to grasp or handle, lifting and carrying 10 lbs, raising arms over head |
| Upper Extremity Functional Scale (1997) | UEFS Nontrauma | Chronic work-related upper extremity disorders; CTS | Multiregion PRO | P | Psychometric evaluation and validation versus work status; physical findings (grip, pinch, Phalen test); duration of symptoms |
| Upper Limb Functional Index (2006) | ULFI Combo | Upper limb symptoms inc postoperative, acute postfracture, ligament sprain patients nos | Multiregion PRO | P | Psychometric evaluation and validation versus DASH, UEFS |
| Shoulder | American Shoulder and Elbow Surgeons Assessment (1994) | ASES-S Combo | Impingement syndrome; instability/dislocation; RCT; adhesive capsulitis; hemiarthroplasty; shoulder weakness; humeral fracture; rotator cuff and adhesive capsulitis; status—postsurgery | Multiregion *PRO | P | Psychometric evaluation and validation versus UPenn Shoulder Score; SF-36 |
| Athletic Shoulder Outcome Scoring System (1993) | | | | | |
| Bostrom Shoulder Impairment Scale (1991) | | | | | |

ACJ, acromioclavicular joint; AIMS2, Arthritis Impact Measurement Scale 2; CBO, clinician-based outcome measure; CMCJ, carpometacarpal joint; Combo, combination of traumatic and nontraumatic conditions; CTS, carpal tunnel syndrome; DRF, distal radius fracture; EPS, elbow plica syndrome; GHJ, glenohumeral joint; GROC, Global Rating of Change; IRGL, Impact of Rheumatic diseases on General health and Lifestyle; LCL, lateral collateral ligament; MCPJ, metacarpophalangeal joint; MEPS, Mayo Elbow Performance Score; MUA, manipulation under anesthetic; NE, no initial empirical psychometric evaluation; NHP, Nottingham Health Profile; nos, nonspecified or nonspecific; NV, no validation studies; OA, osteoarthritis; OCD, osteochondritis dissecans; P, confirmed psychometric evaluation; PHF, proximal humerus fracture; PRO, patient-reported outcome measure; RCT, rotator cuff tear; RhA, rheumatoid arthritis; ROM, range of motion; SCJ, sternoclavicular joint; SF-36, 36-Item Short Form Survey, SIP, sickness impact profile; SLAP, superior labral AP tear; TER, total elbow replacement, US, ultrasound; UTS, ulnar tunnel syndrome; VAS, visual analogue scale. * ASES patient-reported component only—selected as clinician based measure rarely used.
### Table 1 (continued)

#### Outcome Measures for Upper Extremity Conditions Classified by Conditions, Instrument, and Initial Psychometric Evaluation

| Region | Outcome Measure | Short Title | Clinical Conditions in Index Evaluation Study | Instrument Classification | Psychometric Evaluation |
|--------|-----------------|-------------|-----------------------------------------------|---------------------------|-------------------------|
|        |                 |             | Condition Type | Conditions | Coverage Type | Level | Evaluation |
|        |                 |             | Shoulder Score (1985) | CMS | Combo Rotator cuff repair; shoulder arthroplasty; adhesive capsulitis; PHF | Region specific | CBO | NE |
|        |                 |             | Trauma Score for acromioclavicular joint separation (1980) | FLEX-SF | nos nos | Region specific | PRO | NE; NV |
|        |                 |             | Functional Conditions and Function | Haryman rotator cuff functional assessment (1991) | Nontrauma Chronic RCT | Condition specific | PRO | NE; NV |
|        |                 |             | Nontrauma | Hospital for Special Surgery Shoulder Assessment (1982) | HSS Shoulder Rating | Region specific | CBO | NE; NV |
|        |                 |             | Region specific | Hospital for Special Surgery Shoulder Rating Score (1990) | HSS Shoulder Rating | Region specific | CBO | NE; NV |
|        |                 |             | Ipsiatlateral clavicle and scapular neck fractures | Trauma | Hospital of the University of Pennsylvania Shoulder Score (1994) | Nontrauma Open or arthroscopic acromioplasty; chronic impingement syndrome | Condition specific | CBO | NE; NV |
|        |                 |             | Region specific | Imatani Acromioclavicular Separation Evaluation System (1975) | nos | Region specific | CBO | NE; NV |
|        |                 |             | Region specific | Hospital of the University of Pennsylvania Shoulder Score (1994) | nos | Region specific | CBO | NE; NV |
|        |                 |             | Region specific | Japanese Orthopaedic Association Shoulder Score (2004) | nos | Region specific | CBO | NE; NV |
|        |                 |             | Region specific | Kerlan Jobe Orthopaedic Clinic Score (2010) | KJOC | Region specific | PRO | P |
|        |                 |             | Region specific | KSS | KSS | Region specific | CBO | P |
|        |                 |             | Region specific | McGinnis and Denton Rating Scale for Scapular Fractures (1989) | nos | Region specific | CBO | NE; NV |
|        |                 |             | Region specific | Melbourne Instability Shoulder Scale (2005) | MISS | Region specific | PRO | P |

**Legend:**
- ACJ, acromioclavicular joint; AIMS2, Arthritis Impact Measurement Scale 2; CBO, clinician-based outcome measure; CMCJ, carpometacarpal joint; Combo, combination of traumatic and nontraumatic conditions; CTS, carpal tunnel syndrome; DRF, distal radius fracture; EPS, elbow plica syndrome; GHJ, glenohumeral joint; GROC, Global Rating of Change; IRSL, Impact of Rheumatic diseases on General health and Lifestyle; LCL, lateral collateral ligament; MCPJ, metacarpophalangeal joint; MEPS, Mayo Elbow Performance Score; MUA, manipulation under anesthetic; NE, no initial empirical psychometric evaluation; NHP, Nottingham Health Profile; nos, nonspecified or nonspecific; NV, no validation studies; OA, osteoarthritis; OCD, osteochondritis dissecans; P, confirmed psychometric evaluation; PHF, proximal humerus fracture; PRO, patient-reported outcome measure; RCT, rotator cuff tear; RhA, rheumatoid arthritis; ROM, range of motion; SCJ, sternoclavicular joint; SF-36, 36-Item Short Form Survey; SIP, sickness impact profile; SLAP, superior labral AP tear; TER, total elbow replacement, US, ultrasound; UTS, ulnar tunnel syndrome; VAS, visual analogue scale.
- a ASES patient-reported component only—selected as clinician based measure rarely used.
Table 1 (continued)

| Region | Outcome Measure | Short Title | Clinical Conditions in Index Evaluation Study | Instrument Classification | Psychometric Evaluation |
|--------|-----------------|-------------|-----------------------------------------------|---------------------------|-------------------------|
|        |                 |             | Condition Type | Conditions                                      | Coverage | Type | Level  | Evaluation               |
|        |                 |             |                  | Condition specific | Region specific PRO | P | Psychometric evaluation and validation versus CMS, SPADI, DASH |
|        | Modified Rowe Shoulder Score (2005) | MSQ | Combo | Type II SLAP lesion | CBO | NE |
|        | Munich Shoulder Questionnaire (2012) | NCS | Trauma | PHF | CBO | NE; NV |
|        | Neer Shoulder Score (1970) | NCS | Trauma | SCJ/ACJ injuries; clavicle fractures | CBO | NE; NV |
|        | Oxford Instability Score (1999) | OSS | Nontrauma | Degenerative/inflammatory conditions; impingement ± RCT; calcified rotator cuff deposits; primary or secondary OA; inflammatory arthritis; adhesive capsulitis | Region specific PRO | P | Psychometric evaluation and validation versus SF-36; HAQ; CMS |
|        | Oxford Shoulder Score (1996) | PSS | Combo | Impingement; tendonitis; RCT; instability; adhesive capsulitis; PHF; ACJ/GHJ OA | Region specific PRO | P | Psychometric evaluation and validation versus ROM; muscle Force |
|        | Postfunctional rating for long-head biceps tendinitis (1989) | OSS | Nontrauma | Primary bicipital tendinitis | Condition specific CBO | NE; NV |
|        | Rockwood Score for Sternoclavicular Joint Arthritis (1997) | OSS | Nontrauma | SCJ OA | Condition specific CBO | NE; NV |
|        | Rotator Cuff Quality-of-Life Measure (2000) | OSS | RCT (all causes) | Condition specific PRO | P | Psychometric evaluation and validation versus Functional Shoulder Elevation Test; SF-36; ASES |
|        | Rowe Shoulder Score (1978) | OSS | Trauma | Shoulder dislocation; Bankart procedures for shoulder instability | Condition specific CBO | NE |
|        | Shoulder Activity Level—Rating Scale (2005) | OSS | Combo | RCT; shoulder pain; instability; impingement; adhesive capsulitis; other | Region specific PRO | P | Psychometric evaluation and validation versus SST; age; knee activity rating scale; self-reported shoulder activity |

ACJ, acromioclavicular joint; AIMS2, Arthritis Impact Measurement Scale 2; CBO, clinician-based outcome measure; CMCJ, carpometacarpal joint; Combo, combination of traumatic and nontraumatic conditions; CTS, carpal tunnel syndrome; DRF, distal radius fracture; EPS, elbow plica syndrome; GHJ, glenohumeral joint; GROC, Global Rating of Change; IRGL, Impact of Rheumatic diseases on General health and Lifestyle; LCL, lateral collateral ligament; MCPJ, metacarpophalangeal joint; MEPS, Mayo Elbow Performance Score; MUA, manipulation under anesthetic; NE, no initial empirical psychometric evaluation; NHP, Nottingham Health Profile; nos, nonspecified or nonspecific; NV, no validation studies; OA, osteoarthritis; OCD, osteochondritis dissecans; P, confirmed psychometric evaluation; PHF, proximal humerus fracture; PRO, patient-reported outcome measure; RCT, rotator cuff tear; RhA, rheumatoid arthritis; ROM, range of motion; SCJ, sternoclavicular joint; SF-36, 36-Item Short Form Survey, SIP, sickness impact profile; SLAP, superior labral AP tear; TER, total elbow replacement, US, ultrasound; UTS, ulnar tunnel syndrome; VAS, visual analogue scale.

* ASES patient-reported component only—selected as clinician based measure rarely used.
| Region | Outcome Measure | Short Title | Condition Type | Conditions | Instrument Classification | Psychometric Evaluation |
|--------|-----------------|-------------|----------------|------------|--------------------------|-------------------------|
|        | Shoulder Disability Questionnaire (Dutch) (2000) | SDQ-NL | Combo | Shoulder soft tissue disorders (all causes)—electrotherapy/US plus exercise therapy for RCT | Region specific | PRO | P | Psychometric evaluation for responsiveness |
|        | Shoulder Disability Questionnaire (United Kingdom) (1994) | SDQ-UK | nos | Community and primary care subjects with shoulder pain | Region specific | PRO | P | Psychometric evaluation and validation versus ROM; shoulder power |
|        | Shoulder Function Assessment Scale (1996) | SFAS | Nontrauma | RhA | Condition specific | CBO | P | Psychometric evaluation and validation versus VAS; subjective shoulder function; objective shoulder function in 7 daily activities; radiological shoulder destruction |
|        | Shoulder Function Index (2015) | SFInX | Trauma | PHF | Condition specific | CBO | P | Psychometric evaluation and validation inc Rasch analysis |
|        | Shoulder Pain and Disability Index (1991) | SPADI | nos | Shoulder pain | Region specific | PRO | P | Psychometric evaluation and validation versus ROM |
|        | Shoulder Rating Questionnaire (1994) | SRQ | Nontrauma | Impingement syndrome; glenohumeral instability; complete RCT; GHJ OA; adhesive capsulitis; ACJ OA | Region specific | PRO | P | Psychometric evaluation and validation versus AIMS-2; single Q assessing satisfaction in each domain |
|        | Shoulder Severity Index (1987) | SSI | Nontrauma | Shoulder pain; chronic shoulder disability nos nos | Region specific | PRO | NE | |
|        | Simple Shoulder Test (1993) | SST | nos | Shoulder pain | Region specific | PRO | NE | |
|        | Single Assessment Numeric Evaluation Rating (1999) | SANE | Combo | Postoperative following shoulder dislocation; chronic recurrent subluxations; ACJ separations | Region specific | PRO | P | Psychometric evaluation and validation versus Rowe, ASES |
|        | Stanmore Percentage of Normal Shoulder Assessment (2012) | SPONSA | Combo | Shoulder OA/RhA; revision arthroplasty; subacromial impingement; instability; RCT; nonunion of fracture; adhesive capsulitis | Region specific | PRO | P | Psychometric evaluation and validation versus CMS, OSS |
|        | Subjective Shoulder Rating Scale (1997) | SSRS | Combo | Anterior shoulder reconstructions; subacromial decompressions open and arthroscopic; MUA | Region specific | PRO | P | Psychometric evaluation and validation versus CMS; four-point verbal rating scale |
|        | Subjective Shoulder Value (2007) | SSV | Combo | Rotator cuff repair; shoulder arthroplasty; stabilization for recurrent anterior instability | Region specific | PRO | P | Psychometric evaluation and validation versus CMS |

ACJ, acromioclavicular joint; AIMS2, Arthritis Impact Measurement Scale 2; CBO, clinician-based outcome measure; CMCJ, carpometacarpal joint; Combo, combination of traumatic and nontraumatic conditions; CTS, carpal tunnel syndrome; DRF, distal radius fracture; EPS, elbow plica syndrome; GHJ, glenohumeral joint; GROC, Global Rating of Change; IRGL, Impact of Rheumatic diseases on General health and Lifestyle; LCL, lateral collateral ligament; MCPJ, metacarpophalangeal joint; MEPS, Mayo Elbow Performance Score; MUA, manipulation under anesthetic; NE, no initial empirical psychometric evaluation; NHP, Nottingham Health Profile; nos, nonspecified or nonspecific; NV, no validation studies; OA, osteoarthritis; OCD, osteochondritis dissecans; P, confirmed psychometric evaluation; PHF, proximal humerus fracture; PRO, patient-reported outcome measure; RCT, rotator cuff tear; RhA, rheumatoid arthritis; ROM, range of motion; SCJ, sternoclavicular joint; SF-36, 36-Item Short Form Survey, SIP, sickness impact profile; SLAP, superior labral AP tear; TER, total elbow replacement; US, ultrasound; UTS, ulnar tunnel syndrome; VAS, visual analogue scale.

* ASES patient-reported component only—selected as clinician based measure rarely used.
Table 1 (continued)

| Region | Outcome Measure | Short Title | Clinical Conditions in Index Evaluation Study | Instrument Classification | Psychometric Evaluation |
|--------|-----------------|-------------|---------------------------------------------|---------------------------|-------------------------|
|        |                 |             | Condition Type | Conditions | Coverage Type | Level | Evaluation |
| Swanson Shoulder Score (1989) | Combo | RhA; OA; posttraumatic lesions | Region specific | CBO | NE; NV |
| Thorling Subjective Rating for Subacromial Decompression (1985) | Nontrauma | Subacromial decompression (acromioplasty) for shoulder impingement | Condition specific | CBO | NE; NV |
| UCLA End-Result Score (1986) | Combo | RCT | Condition specific | CBO | NE; NV |
| UCLA Shoulder Score (1981) | UCLA Shoulder | OA: osteonecrosis posttrauma; pseudarthrosis posttrauma; RhA; trauma | Region specific | CBO | NE; NV |
| Walch-Duplay Shoulder Instability Score (1987) | Combo | Anterior shoulder instability | Condition specific | CBO | NE |
| Watson Shoulder Score (1985) | Nontrauma | Chronic RCT | Condition specific | CBO | NE; NV |
| Western Ontario Osteoarthritis of the Shoulder Index (2001) | WOOS | Shoulder OA undergoing hemiarthroplasty or total shoulder arthroplasty | Condition specific | PRO | P |
| Western Ontario Rotator Cuff Index (1998) | WORCI | RCT | Condition specific | PRO | P |
| Western Ontario Shoulder Instability Index (1998) | WOSI | Shoulder instability | Condition specific | PRO | P |
| Wolfgang Criteria for rating results of rotator cuff surgical repair (1974) | Combo | RCT | Condition specific | CBO | NE; NV |
| Elbow | American Shoulder and Elbow Surgeons Assessment-Elbow (1999) | ASES-E | nos | nos | Multiregion | CBO | NE |
| Bishop Rating System (1989) | Combo | Cubital tunnel syndrome— anterior intermuscular transfer of ulnar nerve | Condition specific | CBO | NE; NV |
| Broberg and Morrey Elbow Scale (1986) | BMS | Trauma | Region specific | CBO | NE; NV |
| Conway Scoring System (1992) | Trauma | Medial instability (ulnar collateral ligament reconstruction) | Condition specific | CBO | NE; NV |
| Elbow Function Scale (1984) | Trauma | Olecranon fracture—displaced—tension band wiring | Condition specific | CBO | NE; NV |

ACJ, acromioclavicular joint; AIMS2, Arthritis Impact Measurement Scale 2; CBO, clinician-based outcome measure; CMCJ, carpometacarpal joint; Combo, combination of traumatic and nontraumatic conditions; CTS, carpal tunnel syndrome; DRF, distal radius fracture; EPS, elbow plica syndrome; GHJ, glenohumeral joint; GROC, Global Rating of Change; IRGL, Impact of Rheumatic diseases on General health and Lifestyle; LCL, lateral collateral ligament; MCPJ, metacarpophalangeal joint; MEPS, Mayo Elbow Performance Score; MUA, manipulation under anesthetic; NE, no initial empirical psychometric evaluation; NHP, Nottingham Health Profile; nos, nonspecified or nonspecific; NV, no validation studies; OA, osteoarthritis; OCD, osteochondritis dissecans; P, confirmed psychometric evaluation; PHF, proximal humerus fracture; PRO, patient-reported outcome measure; RCT, rotator cuff tear; RhA, rheumatoid arthritis; ROM, range of motion; SCJ, sternoclavicular joint; SF-36, 36-Item Short Form Survey, SIP, sickness impact profile; SLAP, superior labral AP tear; TER, total elbow replacement, US, ultrasound; UTS, ulnar tunnel syndrome;VAS, visual analogue scale.

a ASES patient-reported component only—selected as clinician based measure rarely used.
| Region | Outcome Measure | Short Title | Classical Conditions in Index Evaluation Study | Instrument Classification | Psychometric Evaluation |
|--------|-----------------|-------------|-----------------------------------------------|---------------------------|-------------------------|
|        |                 |             | Condition Type | Conditions | Coverage Type | Level | Evaluation |
| Elbow Functional Assessment Scale (1999) | EFAS | Nontrauma | RHA | Condition specific | CBO | P | Psychometric evaluation and validation versus HSS assessment scale; HSS total elbow scoring system; MEPI |
| Elbow Self-Assessment Score (2015) | ESAS | Combo | Fracture distal humerus, olecranon, proximal forearm, radial head; dislocation; bursitis; distal biceps rupture; TER; OA; OC; EPS; LCL lesion; reconstruction; epicondylitis; nerve lesion; UTS | Region specific | CBO | P | Psychometric testing and validation versus BMS, PREE, MEPS, OES, QuickDASH |
| Ewald Elbow Scale (1975) | nos | nos | Supracondylar fracture (displaced)—pinning | Region specific | CBO | NE |
| Flynn Criteria (1974) | Trauma | Total elbow replacement; RHA; posttraumatic OA; juvenile RHA | Condition specific | CBO | NE |
| Hospital for Special Surgery Assessment Scale (1980) | HSS Assessment Scale | Combo | Failed total elbow arthroplasty—infection, periprosthetic fracture, recurrent dislocation | Region specific | PRO | NE |
| Hospital for Special Surgery Total Elbow Scoring System (1990) | HSS2 Total Elbow Scoring System | nos | Elbow conditions nos | Region specific | PRO | NE; NV |
| Japanese Orthopaedic Association Elbow Evaluation Score (1992) | nos | nos | Condition specific | CBO | NE |
| Jupiter Functional Rating (1985) | Trauma | Distal humerus fractures—intercondylar fractures | Condition specific | PRO | NE; NV |
| Khalfayan Score (1992) | Trauma | Radial head fractures | Condition specific | CBO | NE; NV |
| Liverpool Elbow Score (2004) | LES | Combo | RHA; OA; posttraumatic OA; TER inc revisions; tennis, Golfer elbow; loose body; OCD; posterior impingement; synovial chondromatosis; ulnar nerve problems | Region specific | CBO | P | Psychometric evaluation and validation versus DASH, NHP, SF-12 |
| Mayo Elbow Performance Index and modifications (1992) | MEPI | Combo | RHA—semiconstrained elbow arthroplasty; elbow fracture-dislocation; coronoid process fracture | Region specific | CBO | NE |
| Modified Bishop Scale (1997) | Combo | Ulnar nerve decompression—transposition —Z lengthening flexor pronator mass | Condition specific | CBO | NE; NV |
| Neviaser Criteria (1977) | Trauma | Elbow dislocation | Condition specific | CBO | NE |

(continued)

ACJ, acromioclavicular joint; AIMS2, Arthritis Impact Measurement Scale 2; CBO, clinician-based outcome measure; CMCJ, carpometacarpal joint; Combo, combination of traumatic and nontraumatic conditions; CTS, carpal tunnel syndrome; DRF, distal radius fracture; EPS, elbow plica syndrome; GHJ, glenohumeral joint; GROC, Global Rating of Change; IRGL, Impact of Rheumatic diseases on General health and Lifestyle; LCL, lateral collateral ligament; MCPJ, metacarpophalangeal joint; MEPS, Mayo Elbow Performance Score; MUA, manipulation under anesthesia; NE, no initial empirical psychometric evaluation; NHP, Nottingham Health Profile; nos, nonspecified or nonspecific; NV, no validation studies; OA, osteoarthritis; OCD, osteochondritis dissecans; P, confirmed psychometric evaluation; PHF, proximal humerus fracture; PRO, patient-reported outcome measure; RCT, rotator cuff tear; RHA, rheumatoid arthritis; ROM, range of motion; SCJ, sternoclavicular joint; SF-36, 36-Item Short Form Survey, SIP, sickness impact profile; SLAP, superior labral AP tear; TER, total elbow replacement; US, ultrasound; UTS, ulnar tunnel syndrome; VAS, visual analogue scale.

a ASES patient-reported component only—selected as clinician based measure rarely used.
| Region | Outcome Measure | Short Title | Condition Type | Conditions | Coverage Type | Level Evaluation |
|--------|-----------------|-------------|----------------|------------|---------------|-----------------|
| **Wrist/hand** | **6-item Carpal Tunnel Syndrome Symptom Scale (2009)** | 6-CTS | Nontrauma | CTS | Condition specific | PRO P |
| **Arm** | **ABILHAND Manual Ability Measure (1998)** | ABILHAND | Nontrauma | RhA—wrist arthrosis | Condition specific | PRO P |
| **Arm** | **Alderson-McGill Hand Function Questionnaire (1999)** | AMHFQ | Nontrauma | CTS | Condition specific | PRO P |
| **Arm** | **Arab Hand Function Index (2004)** | AHFI | Nontrauma | RhA | Condition specific | PRO P |

ACJ, acromioclavicular joint; AIMS2, Arthritis Impact Measurement Scale 2; CBO, clinician-based outcome measure; CMCJ, carpometacarpal joint; Combo, combination of traumatic and nontraumatic conditions; CTS, carpal tunnel syndrome; DRF, distal radius fracture; EPS, elbow plica syndrome; GHJ, glenohumeral joint; GROC, Global Rating of Change; IRGL, Impact of Rheumatic diseases on General health and Lifestyle; LCL, lateral collateral ligament; MCPJ, metacarpophalangeal joint; MEPS, Mayo Elbow Performance Score; MUA, manipulation under anesthetic; NE, no initial empirical psychometric evaluation; NHP, Nottingham Health Profile; nos, nonspecified or nonspecific; NV, no validation studies; OA, osteoarthritis; OCD, osteochondritis dissecans; P, confirmed psychometric evaluation; PHF, proximal humerus fracture; PRO, patient-reported outcome measure; RCT, rotator cuff tear; RhA, rheumatoid arthritis; ROM, range of motion; SCJ, sternoclavicular joint; SF-36, 36-item Short Form Survey, SIP, sickness impact profile; SLAP, superior labral AP tear; TER, total elbow replacement, US, ultrasound; UTS, ulnar tunnel syndrome; VAS, visual analogue scale.

a ASES patient-reported component only—selected as clinician based measure rarely used.
### Table 1 (continued)

| Region                  | Outcome Measure                                      | Short Title                      | Clinical Conditions in Index Evaluation Study | Instrument Classification | Psychometric Evaluation |
|-------------------------|------------------------------------------------------|----------------------------------|-----------------------------------------------|----------------------------|-------------------------|
|                         |                                                      |                                  | Condition Type                                 | Conditions                 | Coverage Level Evaluation |
| Arthritis Hand Function Test (1991) | AHFT                                                  | Nontrauma RhA                     | Condition specific                            | CBO                        | P                       |
|                         |                                                      |                                  |                                                | Psychometric evaluation    | versus                    |
|                         |                                                      |                                  |                                                | Australian Hand Impact Measurement Scale 2 | Jebsen Hand Function Test |
| Australian/Canadian Osteoarthritis Hand Index (2002) | AUSCAN                                               | Nontrauma OA                      | Condition specific                            | PRO                        | NE                      |
|                         |                                                      |                                  |                                                | Psychometric evaluation    | validation                |
| Boston Carpal Tunnel Questionnaire (1993) | BCTQ                                                  | Nontrauma CTS                     | Condition specific                            | PRO                        | P                       |
|                         |                                                      |                                  |                                                | Psychometric evaluation    | and validation            |
|                         |                                                      |                                  |                                                | grip and pinch strength, 2-point discrimination, Semmes-Weinstein monofilament test |
| Buck-Gramcko and Lohman Evaluation for Total Wrist Function (1985) | nos                                                   | Compression wrist arthrodesis     | Region specific                             | CBO                        | NE; NV                  |
|                         |                                                      |                                  |                                                | Psychometric evaluation    | and validation             |
|                         |                                                      |                                  |                                                | versus                     |
| Castaing Score (1964)   | Trauma                                               | DRF                              | Condition specific                            | CBO                        | NE; NV                  |
|                         |                                                      |                                  |                                                | Psychometric evaluation    | validation                |
|                         |                                                      |                                  |                                                | versus                     |
|                         |                                                      |                                  |                                                | grip and pinch strength, 2-point discrimination, Semmes-Weinstein monofilament test |
| Clawson Functional Index (1971) | Nontrauma                                             | RhA                              | Condition specific                            | CBO                        | NE; NV                  |
|                         |                                                      |                                  |                                                | Psychometric evaluation    | validation                |
|                         |                                                      |                                  |                                                | versus                     |
|                         |                                                      |                                  |                                                | grip and pinch strength, 2-point discrimination, Semmes-Weinstein monofilament test |
| Colville Quality of Life Hand Questionnaire (1999) | Nontrauma                                             | OA—trapeziectomy; RhA—Swanson MCPJ arthroplasty | Region specific                             | CBO                        | NE; NV                  |
|                         |                                                      |                                  |                                                | Psychometric evaluation    | validation                |
|                         |                                                      |                                  |                                                | versus                     |
|                         |                                                      |                                  |                                                | grip and pinch strength, 2-point discrimination, Semmes-Weinstein monofilament test |
| Crawford Classification (1984) | Trauma                                                | Mallet finger                     | Condition specific                            | CBO                        | NE                      |
|                         |                                                      |                                  |                                                | Psychometric evaluation    | validation                |
|                         |                                                      |                                  |                                                | versus                     |
|                         |                                                      |                                  |                                                | grip and pinch strength, 2-point discrimination, Semmes-Weinstein monofilament test |
| Fernandez Point-Score System (1988) | Trauma                                               | DRF—malunion—radial osteotomy/Bower arthroplasty | Condition specific                             | CBO                        | NE; NV                  |
|                         |                                                      |                                  |                                                | Psychometric evaluation    | validation                |
|                         |                                                      |                                  |                                                | versus                     |
|                         |                                                      |                                  |                                                | grip and pinch strength, 2-point discrimination, Semmes-Weinstein monofilament test |
| Fernandez Scale (1982)  | Trauma                                               | DRF—Posttraumatic correction wrist deformity inc osteotomy, bone grafting, internal fixation | Condition specific                             | CBO                        | NE; NV                  |
|                         |                                                      |                                  |                                                | Psychometric evaluation    | validation                |
|                         |                                                      |                                  |                                                | versus                     |
|                         |                                                      |                                  |                                                | grip and pinch strength, 2-point discrimination, Semmes-Weinstein monofilament test |
| Forearm Symptom Severity Scale (1998) | Trauma                                               | DRF                              | Condition specific                            | CBO                        | P                      |
|                         |                                                      |                                  |                                                | Psychometric evaluation    | validation versus         |
|                         |                                                      |                                  |                                                | Gartland-Werley Scoring system; patient’s and investigator’s subjective characterization of function |
| Functional Index (1984) | Trauma                                               | DRF                              | Condition specific                            | CBO                        | P                      |
|                         |                                                      |                                  |                                                | Psychometric evaluation    | validation versus         |
|                         |                                                      |                                  |                                                | Gartland-Werley Scoring system; patient’s and investigator’s subjective characterization of function |
| Functional Index for Arthropathies of the Hand (1995) | FIHOA                                                | Nontrauma Hand OA—digital or trapeziometacarpal OA—inactive hand OA | Condition specific                             | CBO                        | P                      |
|                         |                                                      |                                  |                                                | Psychometric evaluation    | validation versus         |
|                         |                                                      |                                  |                                                | Gartland-Werley Scoring system; patient’s and investigator’s subjective characterization of function |
| Gartland and Werley Scoring System (1951) | Trauma                                               | DRF                              | Condition specific                            | CBO                        | NE                     |
|                         |                                                      |                                  |                                                | Psychometric evaluation    | validation versus         |
|                         |                                                      |                                  |                                                | Gartland-Werley Scoring system; patient’s and investigator’s subjective characterization of function |

ACJ, acromioclavicular joint; AIMS2, Arthritis Impact Measurement Scale 2; CBO, clinician-based outcome measure; CMCJ, carpometacarpal joint; Combo, combination of traumatic and nontraumatic conditions; CTS, carpal tunnel syndrome; DRF, distal radius fracture; EPS, elbow plica syndrome; GHJ, glenohumeral joint; GROC, Global Rating of Change; IRGL, Impact of Rheumatic diseases on General health and Lifestyle; LCL, lateral collateral ligament; MCPJ, metacarpophalangeal joint; MEPS, Mayo Elbow Performance Score; MUA, manipulation under anesthetic; NE, no initial empirical psychometric evaluation; NHP, Nottingham Health Profile; nos, nonspecified or nonspecific; NV, no validation studies; OA, osteoarthritis; OCD, osteochondritis dissecans; P, confirmed psychometric evaluation; PHF, proximal humerus fracture; PRO, patient-reported outcome measure; RCT, rotator cuff tear; RhA, rheumatoid arthritis; ROM, range of motion; SCJ, sternoclavicular joint; SF-36, 36-Item Short Form Survey, SIP, sickness impact profile; SLAP, superior labral AP tear; TER, total elbow replacement, US, ultrasound; UTS, ulnar tunnel syndrome; VAS, visual analogue scale.

* ASES patient-reported component only—selected as clinician based measure rarely used.
| Region | Outcome Measure | Short Title | Condition Type | Conditions | Instrument Classification | Psychometric Evaluation |
|--------|-----------------|-------------|----------------|------------|--------------------------|------------------------|
|        | Glickel Clinical Grading system (1984) | Nontrauma | Chronic carpal instability | Condition specific | CBO | NE |
|        | Grace and Eversmann Rating (1980) | Trauma | Forearm fractures | Condition specific | CBO | NE |
|        | Green and O’Brien Scoring System (1978) | Trauma | Carpal dislocation—open | Condition specific | CBO | NE; NV |
|        | Hand Function Score (1998) | HFS | Trauma | Hand trauma—rehabilitation nos | Region specific | PRO | NE |
|        | Hand Function Sort (1996) | nos nos | nos nos | Region specific | PRO | NE |
|        | Hand Injury Severity Score (1996) | HFI | Nontrauma | RhA | Condition specific | CBO | NE |
|        | Hospital for Special Surgery Wrist Scoring System (1990) | Non-Trauma | RhA—wrist—total wrist arthroplasty | Condition specific | CBO | NE; NV |
|        | Jebsen-Taylor Hand Function Test (1969) | nos nos | nos nos | Region specific | CBO | NE |
|        | Kapandji Index (1987) | Non-Trauma | RhA | Condition specific | CBO | NE |
|        | Lambert and Clayton Wrist Score (1980) | Non-Trauma | RhA—wrist—total wrist arthroplasty | Condition specific | CBO | NE; NV |
|        | MacBain Hand Function Test (1970) | Nontrauma | RhA | Condition specific | CBO | NE; NV |
|        | Manual Ability Measure—16; -36 (2005) | MAM-16; MAM-36 | Combo | RhA; OA; CTS; median nerve neuritis; tenosynovitis; traumatic injuries inc fractures; open wounds; crush | Region specific | PRO | P |
|        | Martini Score (1999) | Trauma | DRF | Condition specific | CBO | NE; NV |
|        | Measure of Activity Performance of the Hand (2010) | Nontrauma | RhA | Condition specific | PRO | P |
|        | Michigan Hand Outcomes Questionnaire (1998) | Combo | Hand disorders—hand injuries; RhA; CTS | Region specific | PRO | P |
|        | Milliken Activities of Daily Living Scale and modifications (1988) | Trauma | Simple and complex upper limb fractures; soft tissue injuries (tendon/nerve lacerations/repairs); crush; amputations; replantation | Region specific | PRO | P |

ACJ, acromioclavicular joint; AIMS2, Arthritis Impact Measurement Scale 2; CBO, clinician-based outcome measure; CMCJ, carpometacarpal joint; Combo, combination of traumatic and nontraumatic conditions; CTS, carpal tunnel syndrome; DRF, distal radius fracture; EPS, elbow plica syndrome; GHJ, glenohumeral joint; GROC, Global Rating of Change; IRGL, Impact of Rheumatic diseases on General health and Lifestyle; LCL, lateral collateral ligament; MCPJ, metacarpophalangeal joint; MEPS, Mayo Elbow Performance Score; MUA, manipulation under anesthetic; NE, no initial empirical psychometric evaluation; NHP, Nottingham Health Profile; nos, nonspecified or nonspecific; NV, no validation studies; OA, osteoarthritis; OCD, osteochondritis dissecans; P, confirmed psychometric evaluation; PHF, proximal humerus fracture; PRO, patient-reported outcome measure; RCT, rotator cuff tear; RhA, rheumatoid arthritis; ROM, range of motion; SCJ, sternoclavicular joint; SF-36, 36-Item Short Form Survey; SIP, sickness impact profile; SLAP, superior labral AP tear; TER, total elbow replacement; US, ultrasound; UTS, ulnar tunnel syndrome; VAS, visual analogue scale.

a ASES patient-reported component only—selected as clinician based measure rarely used.
| Region                        | Outcome Measure                                      | Short Title | Condition Type | Conditions                                                                 | Clinical Conditions in Index Evaluation Study | Instrument Classification | Psychometric Evaluation |
|-------------------------------|------------------------------------------------------|-------------|----------------|----------------------------------------------------------------------------|-----------------------------------------------|---------------------------|-------------------------|
| Modern Activity               | Subjective Survey of 2007 (2008)                    | MASS07      | nos            | nos                                                                       | Region specific                              | PRO                       | P                       |
| Modified Gartland and         | Werley Scoring System (1975)                         |             | Trauma         | Colles fracture                                                          | Condition specific                            | NE                        |                         |
| Modified Green and O’Brien    | Scoring System (1987)                               |             | Trauma         | Wrist fractures—perilunate fracture dislocations                         | Region specific                              | CBO                       | NE; NV                  |
| Modified Martini Score        | (2004)                                               |             | Trauma         | DRF                                                                       | Condition specific                            | PRO                       | P                       |
| Modified Score for Assessment | Chronic Rheumatic Affections of Hands (2004)        | M-SACRAH    | Nontrauma      | RHA; OA                                                                  | Region specific                              | PRO                       | P                       |
| Munich Wrist Questionnaire    | (2016)                                               | MWQ         | Combo          | DRF—metacarpal, scaphoid, other carpal fractures; TFCC tear; synovitis; SL | Condition specific                            | PRO                       |                         |
| New York Orthopedic Hospital  | Wrist Rating Scale (1991)                            | NYOH        | Trauma         | DNF—external fixation                                                    | Condition specific                            | CBO                       | NE; NV                  |
| Patient Evaluation Measure    | (1995)                                               | PEM         | nos            | Hand surgery nos                                                          | Region specific                              | PRO                       | NE                      |
| Patient-Focused Wrist Outcome | (2003)                                               | PFWO        | Combo          | Wrist disorder/injury nos                                                | Region specific                              | PRO                       | P                       |
| Patient Outcomes of Surgery   | Hand/Arm (2004)                                      | POS Hand/Arm| Nontrauma      | CTS; Dupuytren fasciectomy; joint surgery; tendon surgery; mass excision | Region specific                              | PRO                       | P                       |
| Patient-Rated Wrist/Hand      | Evaluation (2004)                                    | PRWHE       | Trauma         | Wrist/hand fractures; carpal instabilities; OA hand; tendon lacerations;  | Region specific                              | PRO                       | P                       |
| Patient-Rated Wrist Evaluation | (1996)                                               | PRWE        | Trauma         | Scaphoid nonunion; Colles fracture                                        | Region specific                              | PRO                       | P                       |

ACJ, acromioclavicular joint; AIMS2, Arthritis Impact Measurement Scale 2; CBO, clinician-based outcome measure; CMCJ, carpometacarpal joint; Combo, combination of traumatic and nontraumatic conditions; CTS, carpal tunnel syndrome; DRF, distal radius fracture; EPS, elbow plica syndrome; GHJ, glenohumeral joint; GROC, Global Rating of Change; IRGL, Impact of Rheumatic diseases on General health and Lifestyle; LCL, lateral collateral ligament; MCPJ, metacarpophalangeal joint; MEPS, Mayo Elbow Performance Score; MUA, manipulation under anesthetic; NE, no initial empirical psychometric evaluation; NHP, Nottingham Health Profile; nos, nonspecified or nonspecific; NV, no validation studies; OA, osteoarthritis; OCD, osteochondritis dissecans; P, confirmed psychometric evaluation; PHF, proximal humerus fracture; PRO, patient-reported outcome measure; RCT, rotator cuff tear; RHA, rheumatoid arthritis; ROM, range of motion; SCJ, sternoclavicular joint; SF-36, 36-Item Short Form Survey, SIP, sickness impact profile; SLAP, superior labral AP tear; TER, total elbow replacement; US, ultrasound; UTS, ulnar tunnel syndrome; VAS, visual analogue scale.

* ASES patient-reported component only—selected as clinician based measure rarely used.
Table 1 (continued)
Outcome Measures for Upper Extremity Conditions Classified by Conditions, Instrument, and Initial Psychometric Evaluation

| Region               | Outcome Measure                          | Short Title | Clinical Conditions in Index Evaluation Study | Instrument Classification | Psychometric Evaluation |
|----------------------|------------------------------------------|-------------|-----------------------------------------------|---------------------------|------------------------|
|                      |                                          |             | Condition Type                                 | Conditions                | Coverage Type          | Level | Evaluation                                      |
| Revelation Functional Index (1989) | RFI | Nontrauma | RhA | Condition specific | PRO | NE |
| Rheumatoid Hand Functional Disability Scale (1996) | | Nontrauma | RhA | Condition specific | PRO | P |
| Score for Assessment and quantification of Chronic Rheumatic Affections of the Hands (2003) | SACRAH | Nontrauma | RhA; OA | Condition specific | PRO | P |
| Sequential Occupational Dexterity Assessment (1996) | SODA | Nontrauma | RhA—Hand surgery; CTS—release resection ulnar head; tenosynovectomy; articular synovectomy; wrist prosthesis. | Region specific | CBO | P |
| Short Version of the Sequential Occupational Dexterity Assessment (1999) | S-SODA | Nontrauma | RhA | Condition specific | CBO | P |
| Solgaard Functional Score System (1988) | | Trauma | DRF | Condition specific | CBO | NE; NV |
| Sollerman Hand Function Test (1995) | | Combo | RhA; finger amputations; nerve injuries; Dupuytren contracture; shoulder-hand-finger syndromes; posttraumatic hand conditions | Region specific | CBO | P |
| Southampton Dupuytren Scoring System (2014) | SDSS | Nontrauma | Dupuytren contracture | Condition specific | PRO | P |
| Stewart Scores (1984) | | Trauma | DRF—functional cast bracing | Condition specific | CBO | NE; NV |
| Unité Rhumatologique des Affections de la Main scale (2011) | URAM | Nontrauma | Dupuytren disease | Condition specific | PRO | P |
| Wrightington Wrist Function Score (1998) | | Combo | Scapholunate instability—Brunelli procedure | Condition specific | CBO | NE; NV |
| Wrist Outcome Measure (2002) | WOM | Trauma | DRF | Condition specific | CBO | NE; NV |

ACJ, acromioclavicular joint; AIMS2, Arthritis Impact Measurement Scale 2; CBO, clinician-based outcome measure; CMCJ, carpometacarpal joint; Combo, combination of traumatic and nontraumatic conditions; CTS, carpal tunnel syndrome; DRF, distal radius fracture; EPS, elbow plica syndrome; GJU, glenohumeral joint; GROC, Global Rating of Change; IRGL, Impact of Rheumatic diseases on General health and Lifestyle; LCL, lateral collateral ligament; MCPJ, metacarpophalangeal joint; MUEP, Mayo Elbow Performance Score; MUA, manipulation under anesthetic; NE, no initial empirical psychometric evaluation; NHP, Nottingham Health Profile; nos, nonspecified or nonspecific; NV, no validation studies; OA, osteoarthritis; OCD, osteocondritis dissecans; P, confirmed psychometric evaluation; PHF, proximal humerus fracture; PRO, patient-reported outcome measure; RCT, rotator cuff tear; RhA, rheumatoid arthritis; ROM, range of motion; SCJ, sternoclavicular joint; SF-36, 36-Item Short Form Survey, SIP, sickness impact profile; SLAP, superior labral AP tear; TER, total elbow replacement, US, ultrasound; UTS, ulnar tunnel syndrome; VAS, visual analogue scale.

a ASES patient-reported component only—selected as clinician based measure rarely used.
and arm strength, did not correlate with PRO measures of disability. Factors such as social independence appeared to more accurately predict PROs than physician based assessments and even mortality in these patients. Similarly, studies involving distal radius fractures demonstrate depression, anxiety, kinesiophobia, and catastrophic thinking as the most important factors influencing disability and rate of recovery. Despite this growing evidence and the rising demand for robust PRO measurement, there remains a lack of clarity regarding the original development, testing, and quality of PRO measures in the context of upper extremity trauma and disability in this population.

**Objectives**

The primary objective was to identify outcome measures developed for upper extremity conditions, focusing on traumatic injuries, and to classify them by anatomic region, condition type, instrument type, and the psychometric evaluation used in their original development. Secondarily, we aimed to assess the methodological quality of original studies, introducing a PRO measure that incorporated trauma patients in their development. We conclude by highlighting the challenges and solutions encountered in measuring outcomes and disability in this population.

**Methods**

**Data Sources**

A broad search strategy was applied to PubMed (MEDLINE from 1946 to 2016), OVIDSP (EMBASE from 1974 to 2016), CINAHL (from 2006 to 2016), and PsycINFO (from 1806 to 2016) electronic databases on July 1, 2016. Search terms related to “upper limb anatomy,” “outcome measurement,” and demographic parameters were combined with the operator AND (Supplemental Digital Content 1, http://links.lww.com/JG9/A2). No restrictions were set in the search fields, and terms were identified in the title and/or abstract without any limits. Further identification was conducted through an internet search engine (Google) and a contemporary atlas of outcome measures. The review is reported according to the PRISMA statement and registered on the PROSPERO system (No. CRD42016046243) (Appendix 1).

**Study Selection**

Studies involving adult patients experiencing any orthopaedic upper extremity condition involving outcome measurement systems were identified. Abstracts were screened by the lead investigator (P.J.) to (1) generate a comprehensive set of outcome measures and (2) track down the original article introducing the measure plus or minus any development and psychometric evaluation
studies, if available. Psychometric evaluations of PRO measures were taken to include assessments of validity, reliability, responsiveness, interpretability, and acceptability. Eligibility assessment selected only the original publications of PRO measures for qualitative and quantitative synthesis. Measures not recognized as multidomain outcome measurement systems, such as those focusing on clinimetric features alone (eg, range of motion, pathoanatomic or radiological grading and classification, and clinical examination tests), single health components (eg, pain, depression, and return to activity), broad diagnostic groups (eg, osteoarthritis and tumor classifications), and health behavior scales, were excluded along with articles not published in English.

**Data Extraction and Data Synthesis**

Data were extracted, synthesized, and recorded using an electronic database (Microsoft Excel, v15.33). Outcome measures were classified by anatomic region, conditions assessed (ie, broad etiology and specific diagnoses), instrument characteristics (ie, coverage and type), and initial level of psychometric evaluation. Measures combining patient-reported and clinician-based components were classified as the latter by default, unless one or the other was more popularly used in the literature. Initial characterization of psychometric evaluation was based on details of validation (construct validity). If none existed, measures were classified with “no initial empirical psychometric evaluation” or “no initial empirical psychometric evaluation and no validation studies identifiable.”

Quality assessment was conducted using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) criteria and 4-point checklist. This is a well-established standard for evaluating methodological quality, design requirements, and preferred statistical analysis of measurement properties of health-related PRO measures. Only original studies involving patients with trauma conditions in the development and psychometric evaluation of instruments were assessed. Contact with authors was made for clarification of conditions when these were nonspecific. Properties were assessed with the lowest rating within a category taken as the score for the section. In addition, data were extracted for generalizability (ie, population characteristics and sampling procedure) and interpretability but were not rated. The items within PRO measures were also categorized as “best fit” into one of the five health domains by three investigators (P.J., D.R., and S.G.) to calculate the proportion (percentage) of each domain as part of the full score, including any instrument weightings. Discordant judgments were resolved through discussions coordinated by the lead author (P.J.) to reach a consensus between the three investigators. Synthesized data were reported using descriptive statistics and discordant judgments resolved by a discussion among all the authors.

**Results**

A total of 144 outcome measures targeting the upper extremity were identified (Figure 1 and Table 1). The majority focused on the shoulder, wrist, and hand (Figure 2). Fifty-eight percent (n = 83/144) included patients with trauma problems either in combination with other conditions or alone (Figure 3). Seven percent (n = 10/144) required corresponding authors to be contacted to determine conditions investigated because these could not be otherwise identified. Conditions included fractures, dislocations, and soft-tissue injuries (Table 1). Instrument classification revealed the majority as completely or partially clinician based (53%; n = 76/144) (Figure 4) and predominantly condition specific (56%; n = 80/144).
(Figure 5), with trends in instrument type and coverage mapped over time (Figures 6 and 7).

Quality assessment was conducted on 29 original studies (20%; n = 29/144) that included some form of psychometric evaluation of PRO measures when they were first published and involved upper limb trauma patients in their study cohort (Table 2). The majority of studies included in quality assessment were prospective cohort studies. “Test-retest” reliability was assessed more frequently than internal consistency and measurement error, and rating was “poor” to “good” when assessed. Content (face) validity ratings were “good” to “excellent” for almost all measures. Construct validity assessed through testing of hypotheses rated “good to excellent” in two thirds of studies and “poor to fair” in the rest, while structural validity was “poor” in 71% of studies (n = 17/24) when it was assessed. This was primarily due to few studies undertaking factor analysis or item response theory (IRT) analysis, a requisite for higher ratings. The lack of gold-standard measures in this field meant that criterion validity was rarely assessed. Responsiveness was highly variable; although most studies allowed some interpretability through score distribution and change, few conducted analysis of floor-ceiling effects, minimal clinical important difference, or minimal detectable change. Characterization of health domains revealed that the majority of items were related to physical function and symptoms, whereas the relative proportion represented by social (median 14%; range 3%–35%) and psychological aspects (median 14%; range 3%–16.5%) was low. Generalizability assessment revealed low levels of reporting patient consent, percentage of missing items, and study limitations.

**Discussion**

The selection of outcome measures is of paramount importance in conducting high-quality orthopaedic research. Arriving at this choice, particularly among the current assortment of PRO measures, may benefit from an understanding of the methodological quality of their development and relevance to study populations. Over 140 different outcome measures targeted at upper extremity problems were identified, with a substantial number involving trauma conditions in their original cohorts, many of which included distal radius fractures, rotator cuff tears, and shoulder instability. The majority were clinician-based, injury-specific, or procedure-specific instruments and lacked empirical psychometric evaluation in initial development or any identifiable validation studies since introduction. One in five was a PRO measure, involving trauma patients who had undergone initial psychometric evaluation. Methodological quality was deemed acceptable in terms of test-retest reliability, content, and construct validity through hypothesis testing, but was of variable quality and/or lacking in others such
### Table 2

Characterization and Quality Assessment of PRO Measures Involving Patients With Upper Extremity Trauma in the Original Study Cohort

| Outcome Measure | Instrument Items (Score Range) | Physical Function (Activities of Daily Life-Related), % | Psychological, % | Social, % | Pain/ Symptoms, % | Other | Level of Evidence of Index Study(s) | Psychometric Analysis: Property (n) | Outcome Measures |
|-----------------|-------------------------------|------------------------------------------------------|----------------|----------|----------------|-------|---------------------------------|---------------------------------|----------------|
| DASH18          | 30 (0–100)                    | 77                                                   | 3              | 3        | 17             |       | 1b                              | Reliability (407); validity (407); interpretability | VAS for overall problem; VAS for overall pain; VAS of ability to function; VAS of ability to work |
| QuickDASH29     | 11 (0–100)                    | 64                                                   | 18             | 18       |                |       | 1b                              | Reliability (407); validity (407); interpretability | VAS for overall problem; VAS for overall pain; VAS of ability to function; VAS of ability to work |
| UEF20           | 33 (0–99)                     | 100                                                  |                |          |                |       | 3b                              | Reliability (30); validity (79); interpretability | Hand activities of daily living |
| ULFI21          | 25 (0–100)                    | 76                                                   | 4              | 12       | 8              |       | 1b                              | Reliability (64, 32); validity (64); responsiveness (24); interpretability | DASH; UEFs |
| ASES-S22        | 11 (0–100)                    | 45                                                   | 5              | 50       |                |       | 1b                              | Reliability (63); validity (63); responsiveness (63); interpretability | UPenn; SF-36 |
| KJOC23          | 10 (0–100)                    | 50                                                   | 20             | 20       |                |       | 3                               | Reliability (21); validity (282); responsiveness (55); interpretability | DASH; DASH sports/ performing arts module |
| MIS24           | 21 (0–100)                    | 57                                                   | 18             | 25       |                |       | 1b                              | Reliability (22); validity (84); interpretability | SRO; Patient Subjective Rating Scale |
| MSQ25           | 30 (0–314)                    | 62                                                   | 19             | 19       |                |       | 3                               | Validity (56); interpretability | CMS; DASH; SPADI |
| NCS26           | 10 (20–100)                   | 20                                                   | 20             | 70       | 10% (Cosmetic satisfaction) |       | 2b                              | Reliability (70, 50); validity (70); interpretability | CMS; OSS; EQ5D |
| OIS27           | 12 (12–60)                    | 41                                                   | 17             | 25       | 17             |       | 1b                              | Reliability (34, 92); validity (82); responsiveness (15); interpretability | CMS; Rowe; SF-36 |
| PSS28           | 24 (0–100)                    | 54                                                   | 6              | 30       | 10% (Functional satisfaction) |       | 3b                              | Reliability (109, 40); validity (40); responsiveness (109); interpretability | ASES; CMS |
| RC-QOL29        | 34 (0–100)                    | 35                                                   | 15             | 35       | 15             |       | 1b                              | Reliability (30); validity (86); interpretability | Functional shoulder elevation test; SF-36; ASES |
| SAS30           | 5 + 2 (0–20)                  | 100                                                  |                |          |                |       | 1b                              | Reliability (40); validity (42); interpretability | SST; age; knee activity rating scale; self-reported shoulder activity (continued) |

IRT = item response theory; MCID = minimal clinical important difference; MDC = minimal detectable change; PRO = patient-reported outcome; SAS = Shoulder Activity Scale.

a. Including instrument weighting.

b. COSMIN Box G (Cross-cultural validity) was excluded—translations and cultural adaptations out of scope.
### Table 2 (continued)

Characterization and Quality Assessment of PRO Measures Involving Patients With Upper Extremity Trauma in the Original Study Cohort

| Outcome Measure | Instrument Items (Score Range) | Physical Function (Activities of Daily Life-Related), % | Psychological, % | Social, % | Pain/ Symptoms, % | Other | Study Characteristics and Psychometric Evaluation | Level of Evidence of Index Study(s) | Psychometric Analysis: Property (n) | Outcome Measures |
|-----------------|--------------------------------|-------------------------------------------------------|----------------|-----------|-----------------|-------|---------------------------------------------------|----------------------------------|-----------------------------------|------------------|
| SDQ-NL31        | 16 (0–100)                     | 94 (Linked to pain)                                   | 6              |           | 88 (Linked to physical function—daily activities) |       | 1b Reliability (180); responsiveness (180); interpretability | Patient-rated VAS for severity of shoulder pain and VAS chief complaint; 8-point Likert Scale for overall change since baseline; clinical rating VAS symptom severity and VAS mobility restriction |
| SANE32          | 1 (0–100)                      | 100% (Global general rating)                          |               |           |                 | 1b    | Reliability (163); validity (163); interpretability | Rowe; ASES |
| SPONSA33        | 1 (0–100)                      | 100% (Global general rating inc pain and physical function) |               |           |                 | 1b    | Reliability (61); validity (61); responsiveness (61); interpretability | CMS; OSS |
| SSRS34          | 5 (0–100)                      | 35                                                     | 15             | 50        |                 | 1b    | Reliability (200); validity (200); interpretability | CMS; four-point verbal rating scale |
| SSV35           | 1 (0–100)                      | 100% (Global general rating)                          |               |           |                 | 3b    | Reliability (441); validity (441); interpretability | CMS |
| WORC36          | 21 (0–2,100)                   | 47                                                     | 14             | 10        | 29              | 3b    | Reliability (100); validity (97); responsiveness (100); interpretability | UCLA; SF-36; CMS; ASES; DASH; SF-12; ROM |
| WOSI37          | 21 (0–2,100)                   | 24                                                     | 14             | 24        | 38              | 3b    | Reliability (51); validity (47); responsiveness (47); interpretability | DASH; ASES; UCLA; Rowe; CMS; SF-12; global change; ROM |
| OES38           | 12 (0–48)                      | 33                                                     | 16.5           | 16.5      | 33              | 1b    | Reliability (104, 52); validity (104); responsiveness (104); interpretability | MEPS; DASH; SF-36 |
| PREE39          | 20 (0–100)                     | 54                                                     | 13             | 33        |                 | 1b    | Reliability (70); validity (70); interpretability | ASES; SF-36; DASH |
| MAM-16; MAM-3640 | 16 (0–100)                    | 100                                                    |               |           |                 | 1b    | Reliability (115); validity (115)                   | LIFeware Musculoskeletal Form; SF-12 health status |
| MHQ41           | 37 (0–100)                     | 46                                                     | 14             | 14        | 10% (Appearance); 16% (satisfaction in relation to physical function, pain, sensation) | 1b    | Reliability (22); validity (200); interpretability | SF-12 |

IRT = item response theory; MCID = minimal clinical important difference; MDC = minimal detectable change; PRO = patient-reported outcome; SAS = Shoulder Activity Scale.

* Including instrument weighting.

** COSMIN Box G (Cross-cultural validity) was excluded—translations and cultural adaptations out of scope.
| Outcome Measure | Instrument Items (Score Range) | Physical Function (Activities of Daily Life–Related), % | Psychological, % | Social, % | Pain/ Symptoms, % | Other | Study Characteristics and Psychometric Evaluation |
|-----------------|--------------------------------|----------------------------------------------------------|------------------|-----------|------------------|-------|--------------------------------------------------|
| MAS42           | 47 (47–235)                    | 100                                                      |                  |           |                  |       | 1b Reliability (45); validity (37); interpretability |
| MWO43           | 16 (0–100)                     | 63                                                       | 6                | 25        | 6% (Satisfaction) |       | 1b Reliability (100); validity (100); responsiveness (100) |
| PFWO44,45       | 52 (Variable) 95 (Approx)      | 5 (Approx)                                               |                  |           |                  |       | 1b Reliability (50); validity (50, 26); responsiveness (26); interpretability |
| PRWHE46         | 15 (0–100)                     | 54                                                       | 13               | 33        | Plus 10 pts (Appearance) |       | 1b Validity (60); responsiveness (60); interpretability |
| PRWE47          | 15 (0–100)                     | 54                                                       | 13               | 33        |                  |       | 3b Reliability (38); validity (53) SF-36; impairment score (wrist ROM, grip strength; checkers subset) |

COSMIN Checklist and Four-Point Rating Systemb

| Reliability | Content Validity | Construct Validity | Criterion Validity | Responsiveness | Interpretability |
|-------------|------------------|--------------------|--------------------|----------------|------------------|
| A. Internal Consistency | B. Reliability (Test-Retest, Interrater) | C. Measurement Error (Test-Retest, Interrater) | D. Content (Face) Validity | E. Structural Validity | F. Hypothesis Testing | G. Criterion Validity (Concurrent Validity, Predictive Validity) | H. Responsiveness | I. Interpretability |
| DASH18       | 0                | Excellent          | 0                  | Excellent      | Excellent        | 0                  | Excellent              | Score distribution; score change |
| QuickDASH19  | 0                | Excellent          | 0                  | Excellent      | Excellent        | 0                  | Excellent              | Score distribution; score change |
| UEFT20       | 0                | Poor               | 0                  | Poor           | Poor             | 0                  | 0                     | Score distribution; score change |
| ULFI21       | 0                | Good               | Good               | Excellent      | Poor             | Good               | Good                  | Missing item handling; score distribution; score change; floor-ceiling effect; MDC |
| ASES-S22     | 0                | Good               | Good               | Excellent      | Poor             | Good               | Good                  | Score distribution; score change; MDC; MCID |

IRT = item response theory; MCID = minimal clinical important difference; MDC = minimal detectable change; PRO = patient-reported outcome; SAS = Shoulder Activity Scale.

a Including instrument weighting.

b COSMIN Box G (Cross-cultural validity) was excluded—translations and cultural adaptations out of scope.
Table 2 (continued)

Characterization and Quality Assessment of PRO Measures Involving Patients With Upper Extremity Trauma in the Original Study Cohort

| Outcome Measure | IRT Box | A. Internal Consistency | B. Reliability (Test-Retest, Interrater, Intrarater) | C. Measurement Error (Test-Retest, Interrater, Intrarater) | D. Content (Face) Validity | E. Structural Validity | F. Hypothesis Testing | G. Criterion Validity (Concurrent Validity, Predictive Validity) | H. Responsiveness | I. Interpretability |
|----------------|---------|------------------------|------------------------------------------------------|-----------------------------------------------------------|---------------------------|----------------------|---------------------|---------------------------------------------------------------|-----------------|-------------------|
| KJOC23         | 0       | Poor                   | Poor                                                 | Poor                                                      | Excellent                 | Good                 | Fair                | 0                                                                             | Good            | Score distribution; score change                              |
| MISS24         | 0 0     | Poor                   | Poor                                                 | Poor                                                      | Good                      | Poor                 | Fair                | 0                                                                             | 0               | Score distribution; score change; MDC                        |
| MSQ25          | 0 0 0   | Good                   | Poor                                                 | Good                                                      | Poor                      | Good                 | 0                   | Score distribution; floor-ceiling effects                        | Good            | Score distribution; score change                              |
| NCS26          | 0 Fair  | Poor                   | 0                                                    | Good                                                      | Good                      | Poor                 | Good                | 0                                                                             | 0               | Score distribution; floor-ceiling effects; score change; MDC, MCID |
| OIS27          | 0 Fair  | Good                   | 0                                                    | Excellent                                                 | Poor                      | Good                 | 0                   | Good                                                                          | Score distribution; score change                              |
| PSS28          | 0 Good  | Fair                   | Fair                                                 | Fair                                                      | Excellent                 | Poor                 | Fair                | 0                                                                             | Good            | Score distribution; floor-ceiling effects; score change; MDC, MCID |
| RC-QOL29       | 0 0     | Fair                   | 0                                                    | Excellent                                                 | Poor                      | Good                 | 0                   | Score distribution; floor-ceiling effects; score change              | 0               | Score distribution; score change                              |
| SAS30          | 0 0     | Fair                   | 0                                                    | Excellent                                                 | Poor                      | Fair                 | 0                   | Score distribution                                                | Score distribution; score change                              |
| SDQ-NL31       | 0 0 0   | Good                   | 0                                                    | 0                                                         | 0                         | 0                   | Fair                | Score distribution; floor-ceiling effects                           | Score distribution; floor-ceiling effects; score change; MDC, MCID |
| SANE32         | 0 0 0   | Good                   | 0                                                    | Good                                                      | 0                         | Good                 | 0                   | Score distribution; score change                                     | Score distribution; score change                              |
| SPONSA33       | 0 0     | Fair                   | 0                                                    | Good                                                      | 0                         | Good                 | 0                   | Score distribution; score change                                     | Good            | Score distribution; score change                              |
| SSRS34         | 0 0     | Fair                   | 0                                                    | Good                                                      | 0                         | Fair                 | 0                   | Score distribution; score change                                     | Score distribution; score change                              |
| SSV35          | 0 0     | Fair                   | 0                                                    | Good                                                      | Poor                      | Good                 | 0                   | Score distribution; score change                                     | Good            | Score distribution; score change                              |
| WORCI36        | 0 0     | Good                   | 0                                                    | Excellent                                                 | Poor                      | Good                 | 0                   | Good                                                                          | Score change                                              |
| WOSI37         | 0 0     | Good                   | 0                                                    | Excellent                                                 | Poor                      | Fair                 | 0                   | Fair                                                                          | Score change                                              |

IRT = item response theory; MCID = minimal clinical important difference; MDC = minimal detectable change; PRO = patient-reported outcome; SAS = Shoulder Activity Scale.

* Including instrument weighting.

* COSMIN Box G (Cross-cultural validity) was excluded—translations and cultural adaptations out of scope.
as internal consistency, measurement error, responsiveness, and interpretability. This work also demonstrated a transition from pure clinician-based measures to nonvalidated PRO measures to those with some form of validation before introduction (Figure 6). A further trend toward an increase in the development of region-specific instruments over condition-specific measures is observed (Figure 7). This may reflect the importance placed on PRO measurement in modern orthopaedic practice as well as a possible trend toward more general outcome measurements of disability impact at the regional level. Despite these findings, relatively low rates of patient-reported assessment have been observed in the orthopaedic trauma literature, and the drive to develop these instruments does not appear to correlate with their level of utilization in clinical practice.50

Table 2 (continued)

| Characterization and Quality Assessment of PRO Measures Involving Patients With Upper Extremity Trauma in the Original Study Cohort |
|--------------------------------------------------------------------------------------------------------------------------------|
| **COSMIN Checklist and Four-Point Rating System**                                                                                     |
| **Outcome Measure** | **IRT Box** | **A. Internal Consistency** | **B. Reliability (Test-Retest, Interrater, Intra-rater)** | **C. Measurement Error (Test-Retest, Interrater, Intra-rater)** | **D. Content (Face) Validity** | **E. Structural Validity** | **F. Hypothesis Testing** | **G. Criterion Validity (Concurrent Validity, Predictive Validity)** | **H. Responsiveness** | **I. Interpretability** |
|---------------------|-------------|-----------------------------|----------------------------------------------------------|---------------------------------------------------------------|-------------------------------|--------------------------|--------------------------|---------------------------------------------------------------|------------------------|------------------------|
| OES38               | Good        | Excellent                   | Good                                                     | Excellent                                                     | Excellent                    | Good                     | Excellent                | 0                              | Excellent              | Score distribution; floor-ceiling effects; score change |
| PREE39              | 0           | 0                           | Good                                                     | 0                                                             | Excellent                    | Poor                     | Good                     | 0                              | 0                      | Score distribution |
| MAM-16; MAM-3650    | Good        | 0                           | Good                                                     | Excellent                                                     | Good                         | Poor                     | Good                     | 0                              | 0                      | Score distribution |
| MHQ41               | 0           | Poor                        | Poor                                                     | 0                                                             | Excellent                    | Good                     | Good                     | 0                              | 0                      | Score distribution |
| MAS42               | 0           | 0                           | Fair                                                     | 0                                                             | Excellent                    | Poor                     | Fair                     | 0                              | 0                      | Score change          |
| MWQ43               | 0           | Good                        | Good                                                     | 0                                                             | Excellent                    | Poor                     | Good                     | 0                              | Good                  | Score change          |
| PFWO44,45           | 0           | 0                           | Good                                                     | 0                                                             | Good                         | Poor                     | Poor                     | 0                              | Poor                  | Floor-ceiling effect |
| PRWHE46             | 0           | 0                           | 0                                                       | 0                                                             | Excellent                    | 0                        | 0                        | Good                            | 0                      | Score change |
| PRWE47              | 0           | 0                           | Fair                                                     | 0                                                             | Excellent                    | Good                     | Good                     | 0                              | 0                      | Score change |

IRT = item response theory; MCID = minimal clinical important difference; MDC = minimal detectable change; PRO = patient-reported outcome; SAS = Shoulder Activity Scale.

a Including instrument weighting.
b COSMIN Box G (Cross-cultural validity) was excluded—translations and cultural adaptations out of scope.

Limitations

There are some limitations to this work. First, only the original index articles relevant to each outcome measure were within scope. We were interested in understanding the nature of the original developmental work by the investigators and the methodological quality of these measures. Although we located the studies for each instrument, it is recognized that more than one early investigation could lay claim to being part of the initial psychometric evaluation. Furthermore, subsequent studies may have superseded these index evaluations, performing further assessments, including specific trauma populations. Second, although the proportion of psychosocial components within our selection of PRO measurements was judged to be low, it is appreciated that investigators may incorporate other measures to account for psychological and social well being. Third, the identified outcome measurement set is unlikely to be exhaustive, but intuitively any instruments “missed” are more likely clinician-based than...
patient-reported. In this regard, there is an issue of publication bias and its impact on the internal validity of this work. Unpublished studies reporting negative, unfavorable outcomes following instrument testing may exist. The lack of reporting limitations in many of the studies may also reflect a level of reporting bias where there has been a vested interest in instrument promotion. Fourth, we recognize that “all upper extremity trauma conditions” were included as the target category, and findings around methodological quality of measures may vary if the evaluations were performed around specific injuries. Finally, in 7% of authors contacted, inquiries were limited to diagnostic clarification, and no further information was gathered around methodological quality. Thus, it was unclear whether low ratings were down to lack of reporting or actual lack of quality according to COSMIN.

The findings of this review can be considered in light of some of the challenges and solutions in this field. Timing and recruitment of orthopaedic trauma patients for instrument development, testing, and outcome assessment may be logistically difficult because of a variety of clinical and environmental stressors. Measurement should occur at a time when patients are “stable” enough to perform evaluations while being close enough to the date of injury to fully capture the health-related impact. This can be challenging when considering the effects of symptoms (eg, fracture-related pain) and clinical circumstances (eg, fracture immobilization). These issues may be unavoidable but managed best by improved patient and staff education as outcome measurement becomes part of everyday orthopaedic practice.

Responder burden plus inefficient and irrelevant testing are further issues, especially when full and lengthy fixed-length outcome measures are administered in these populations. The risks of incomplete scoring, poor patient experience, “gaming” of the assessment for perceived reward, and difficulties in estimating performance while being “out of action,” alongside the tendency to overestimate one’s level of ability in these situations, are apparent.

Functionality and psychometric properties of instruments primarily developed in chronic conditions may lack adequate coverage and incorporate a set of items too narrow and limited in assessing health impact outside this context. This tendency may be reflected by high floor-ceiling effects during applications in trauma. Furthermore, instruments or groups of instruments...
should adequately cover all relevant health-related domains, including psychosocial factors which are shown to have a dominant influence on disability.

Tailored assessment of patients is an ongoing challenge within outcome measurement in general. It is particularly relevant in orthopaedic trauma where population characteristics are wide ranging. One component involves the assessment of a patient’s baseline status, a particular challenge in trauma situations. Other aspects include the capture of patient experience and emerging concept of patient activation, an individual’s level of involvement in his or her care and the propensity to engage in adaptive health behaviors.28,29,54

One or more of these challenges can be met by the introduction and/or mode of application of established and contemporary outcome measurement systems. First, combinations of region-specific and generic measures with instruments measuring specific factors of relevance, such as depression and pain interference, could be applied to more comprehensively assess patient-focused, health-related outcomes. Instrument choice should depend on the psychometric attributes of the measure, methodological quality in development, and evidence of validation in the target population.49 Generic PRO measures have gained popularity in trauma as they provide a more holistic measurement of health-related outcomes in the multiply injured and medically complex patient, while allowing comparisons between interventions.3 Collaborative efforts are underway to develop standardized outcome sets through a consensus-based selection of generic and specific outcome measures.55

Second, abbreviated versions of well-established scales (eg, QuickDASH) have been developed to improve efficiency and performance while maintaining validity against their full-version counterparts.19 Another contemporary solution involves computerized adaptive tests (CATs). CATs are dynamic tests using computers to administer test items based on the IRT mathematical model.56–58 An IRT-based algorithm allows adaptation to the patient’s last response and administration of relevant subsequent items from a large question bank.58 The Patient-Reported Outcome Measurement Information System (PROMIS) developed by the US National Institute of Health is one of the most commonly used CAT systems.56–58
PROMIS CAT scores range from 0 to 100, with 50 points as US general population mean. They enable capture of physical (eg, physical function and pain interference), mental (eg, anxiety, and depression), and social (eg, social isolation) health domains through modules that can be tailored to the study and population being assessed. Customization, avoiding redundancy, minimizing floor-ceiling effects, and maximizing scoring efficiency and measurement precision are clearly advantageous in the trauma setting. Studies have demonstrated the correlation of CATs with popular fixed-length scales.

In general, computer-based outcome assessment represents a positive paradigm shift, with instruments such as PROMIS CATs being incorporated in outcome measurement software by well-established organizations such as the AO Foundation. It is important to note that PROMIS CATs were originally developed for chronic conditions, and their development as regional measures (eg, PROMIS Upper Extremity Physical Function CAT) and evaluation in traumatic conditions is ongoing. Other outcome measures with adaptive capabilities, but delivered through a paper-based format, include the FLEX-SF shoulder instrument and short-form PROMIS measures.

Some fixed-length scales have also been designed to provide a more relevant, patient-specific assessment by factoring in patient-reported “levels of ability” and “levels of necessity” in relation to various activities. Other instruments have focused on accurate measurement of functional progress and minimizing the discrepancy between what patients report and what they actually do, by clinician-observed grading of enacted activities of daily life in real time. PRO measures such as the PFWO and ULFI are designed to capture recall of preinjury performance and baseline function. The former includes a component that accounts for compensatory mechanisms in performing daily activities. Another strategy in establishing a baseline involves the use of patient proxies, such as family and friends, to aid in recall of preinjury function during the early postinjury phase.

In terms of patient satisfaction with various health domains, the MHQ includes a component measuring satisfaction with appearance, physical function, and symptoms. Patient experience, including satisfaction with care, is often assessed through separate scales, although instruments such as the SRI and MWQ evaluate this aspect in musculoskeletal trauma and wrist/hand injuries, respectively. Early work on patient activation measures has demonstrated a direct correlation with satisfaction among upper extremity conditions and musculoskeletal trauma patients, as well as improved pain relief, mental health, and reduced disability. Further research is necessary to assess correlation with PROs.

This work has systematically reviewed the methodological quality of studies involving PRO measurements in upper extremity trauma on a broad scale. Focused evaluations, using the COSMIN checklist, have been conducted in distal radius fractures; however, the literature in this area is lacking overall. Instrument properties should be defined for the population being tested and not for the PRO instrument itself. In reality, PRO measures have been used throughout orthopaedics in patient groups for which the instrument was not initially developed or psychometrically evaluated. It is unclear whether commonly used instruments can measure all the health-related aspects surrounding upper limb trauma important to the individual. These measures are commonly selected by intuition, clinical culture, and familiarity, with the belief that they are “fit for purpose” and capable of capturing the substantive components of disability experienced by these patients. Ultimately, PRO measurement selection requires careful consideration of the methodological quality, and further research is required to evaluate their psychometric properties in these populations. Reaching a consensus on outcome measurement sets in trauma that are delivered in a standardized fashion will form a more complete, comparable, and interpretable assessment of disability in these populations.

References

1. Suk M, Hanson B, Norvell D, Helfet D: AO Handb Book. Musculoskeletal Outcomes Measures and Instruments. Selection and Assessment Upper Extremity, ed 1. Davos, Switzerland, Thieme, AO Publishing, 2009, vol 1, pp 65-387.
2. Ayers DC, Bozic KJ: The importance of outcome measurement in orthopaedics. Clin Orthop Relat Res 2013;471:3409-3411.
3. World Health Organization, ICF: International Classification of Functioning, Disability and Health. http://www.who.int/classifications/icf/en/. Accessed on June 13, 2017.
4. Bhandari M, Busse JW, Hanson BP, Leece P, Ayeni OR, Schemitsch EH: Psychological distress and quality of life after orthopedic trauma: An observational study. Can J Surg 2008;51:15-22.
5. Kaske S, Lefering R, Trentzsch H, et al: Quality of life two years after severe trauma: A single centre evaluation. Injury 2014;45(suppl 3):S100-S105.
6.Nota SPFT, Bot AGJ, Ring D, Kloen P: Disability and depression after orthopaedic trauma. Injury 2015;46:207-212.
7. Levin PE, MacKenzie EJ, Bosse MJ, Greenhouse PK: Improving outcomes: Understanding the psychosocial aspects of the orthopaedic trauma patient. Instr Course Lect 2014;63:39-48.
8. Vranceanu AM, Bachoura A, Weening A, Vrahais M, Smith RM, Ring D: Psychological factors predict disability and pain intensity after skeletal trauma. J Bone Joint Surg Am 2014;96:e20.
9. Das De S, Vranceanu AM, Ring DC: Contribution of kinesiophobia and catastrophic thinking to upper-extremity-specific disability. J Bone Joint Surg Am 2013;95:76-81.
10. Bot AGJ, Bekkers S, Arnstein PM, Smith RM, Ring D: Opioid use after fracture surgery correlates with pain intensity and satisfaction with pain relief. Clin Orthop Relat Res 2014;472:2542-2549.
11. De Putter CE, Selles RW, Haagsma JA, et al: Health-related quality of life after upper extremity injuries and predictors for suboptimal outcome. Injury 2014;45:1752-1758.
12. Clement ND, Duckworth AD, McQueen MM, Court-Brown CM: The outcome of proximal humeral fractures in the elderly: Predictors of mortality and function. Bone Joint J 2014;96B:970-977.
13. Slobogean GP, Noonan VK, Famanuea A, O'Brien PJ: Does objective shoulder impairment explain patient-reported functional outcome? A study of proximal humerus fractures. J Shoulder Elbow Surg 2011;20:267-272.
14. Ring D, Kadzielski J, Fabian L, Zarukowski D, Malhotra LR, Jupiter JB: Self-reported upper extremity health status correlates with depression. J Bone Joint Surg Am 2006;88:1983-1988.
15. Roh YH, Lee BK, Noh JH, Mohr K, Bowerdy J, Yocum L: The development and validation of a functional assessment tool for the upper extremity in the overhead athlete. Am J Sports Med 2010;38:903-911.
16. Watson L, Story I, Dalziel R, Hoy G, Shinnam A, Woods D: A new clinical outcome measure of glenohumeral joint instability: The MISS questionnaire. J Shoulder Elbow Surg 2005;14:22-30.
17. Schmidutz F, Beier M, Braunstein B, Bogn C, Wiedemann E, Biberthaler P: The Munich Shoulder Questionnaire (MSQ): Development and validation of an effective patient-reported tool for outcome measurement and patient safety in shoulder surgery. Patient Saf Surg 2012;6:9.
18. Charles E, Kumar V, Blacknall J, Edwards S, Geoghegan JM, Manning PA: A validation of the Northbridge Clavicle Score: A clavicle, acromio-clavicular joint and sternoclavicular joint specific patient reported outcome measure. BJ Orthop Proc 2013;9B(supp 1):50.
19. Dawson J, Dool H, Boller I, Fitzpatrick R, Little C, Rees J: The development and validation of a patient-reported questionnaire to assess outcomes of elbow surgery. J Bone Joint Surg Br 2008;90:466-473.
20. Kirkley A, Griffin S, McIntosh 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.
21. Dawson J, Doll H, Boller I, Fitzpatrick R, Little C, Rees J: The development and validation of a patient-reported questionnaire to assess outcomes of elbow surgery. J Bone Joint Surg Br 2008;90:466-473.
22. Kirkley A, Griffin S, McIntosh 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.
23. Dawson J, Nick H, Nick H, Nick H, Nick H: The development and validation of a patient-reported questionnaire to assess outcomes of elbow surgery. J Bone Joint Surg Br 2008;90:466-473.
24. MacDermid JC: Outcome evaluation in patients with elbow pathology: Issues in instrument development and evaluation. J Hand Ther 2001;14:105-114.
25. Chen CC, Granger CV, Peiman CA, Moy OJ, Wald S: Manual Ability Measure (MAM-16): A preliminary report on a new patient-centred and task-orientated outcome measure of hand function. J Hand Surg Br 2005;30:207-216.
26. Chung KC, Pillsbury MS, Walters MR, Hayward RA: Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. J Hand Surg Am 1998;23:575-587.
27. Seaton MK, Grooth GN, Matheson L, Feely C: Reliability and validity of the Milken Activities of Daily Living Scale. J Occup Rehabil 2005;15:343-351.
28. MacDermid JC: Responsiveness of the Disability of the Arm, Shoulder and Hand shoulder assessment. Int J Shoulder Surg 2012;6:9-14.
29. Kohn D, Geyer M: The subjective shoulder rating system. Arch Orthop Trauma Surg 1997;116:324-328.
30. Gilbart MK, Gerber C: Comparison of the subjective shoulder value and the Constant Score. J Shoulder Elbow Surg 2007;16:717-721.
31. Kirkley A, Alvarez C, Griffin S: The development and evaluation of a disease-specific quality of life questionnaire for disorders of the rotator cuff: The Western Ontario Rotator Cuff Index. Clin J Sport Med 2003;13:84-92.
32. Kirkley A, Griffin S, McIntosh 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.
33. Dawson J, Dool H, Boller I, Fitzpatrick R, Little C, Rees J: The development and validation of a patient-reported questionnaire to assess outcomes of elbow surgery. J Bone Joint Surg Br 2008;90:466-473.
34. Chen CC, Granger CV, Peiman CA, Moy OJ, Wald S: Manual Ability Measure (MAM-16): A preliminary report on a new patient-centred and task-orientated outcome measure of hand function. J Hand Surg Br 2005;30:207-216.
35. Chung KC, Pillsbury MS, Walters MR, Hayward RA: Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. J Hand Surg Am 1998;23:575-587.
36. Seaton MK, Grooth GN, Matheson L, Feely C: Reliability and validity of the Milken Activities of Daily Living Scale. J Occup Rehabil 2005;15:343-351.
37. MacDermid JC: Responsiveness of the Disability of the Arm, Shoulder and Hand shoulder assessment. Int J Shoulder Surg 2012;6:9-14.
38. Kohn D, Geyer M: The subjective shoulder rating system. Arch Orthop Trauma Surg 1997;116:324-328.
39. Gilbart MK, Gerber C: Comparison of the subjective shoulder value and the Constant Score. J Shoulder Elbow Surg 2007;16:717-721.
40. Kirkley A, Alvarez C, Griffin S: The development and evaluation of a disease-specific quality of life questionnaire for disorders of the rotator cuff: The Western Ontario Rotator Cuff Index. Clin J Sport Med 2003;13:84-92.
41. Kirkley A, Griffin S, McIntosh 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.
42. Dawson J, Dool H, Boller I, Fitzpatrick R, Little C, Rees J: The development and validation of a patient-reported questionnaire to assess outcomes of elbow surgery. J Bone Joint Surg Br 2008;90:466-473.
43. MacDermid JC: Outcome evaluation in patients with elbow pathology: Issues in instrument development and evaluation. J Hand Ther 2001;14:105-114.
44. Chen CC, Granger CV, Peiman CA, Moy OJ, Wald S: Manual Ability Measure (MAM-16): A preliminary report on a new patient-centred and task-orientated outcome measure of hand function. J Hand Surg Br 2005;30:207-216.
45. Chung KC, Pillsbury MS, Walters MR, Hayward RA: Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. J Hand Surg Am 1998;23:575-587.
46. Seaton MK, Grooth GN, Matheson L, Feely C: Reliability and validity of the Milken Activities of Daily Living Scale. J Occup Rehabil 2005;15:343-351.
47. MacDermid JC: Responsiveness of the Disability of the Arm, Shoulder and Hand shoulder assessment. Int J Shoulder Surg 2012;6:9-14.
48. Kohn D, Geyer M: The subjective shoulder rating system. Arch Orthop Trauma Surg 1997;116:324-328.
49. Gilbart MK, Gerber C: Comparison of the subjective shoulder value and the Constant Score. J Shoulder Elbow Surg 2007;16:717-721.
50. Kirkley A, Alvarez C, Griffin S: The development and evaluation of a disease-specific quality of life questionnaire for disorders of the rotator cuff: The Western Ontario Rotator Cuff Index. Clin J Sport Med 2003;13:84-92.
(DASH) and Patient-Rated Wrist/Hand Evaluation (PRWHE) in evaluating change after hand therapy. *J Hand Ther* 2004;17:18-23.

47. MacDermid JC: Development of a scale for patient rating of wrist pain and disability. *J Hand Ther* 1996;9:178-183.

48. Bhandari M, Petrisor B, Schemitsch E: Outcome measurements in orthopaedics. *Indian J Orthop* 2007;41:32-36.

49. Kleinlugtenbelt YV, Nienhuis RW, Bhandari M, Goslings JC, Poolman RW, Scholtes VAB: Are validated outcome measures used in distal radial fractures truly valid? *Bone Joint Res* 2016;5:153-161.

50. Horwitz DS, Richard RD, Suk M: The reporting of functional outcome instruments in the Journal of Orthopaedic Trauma over a 5-year period. *J Orthop Trauma* 2014;28:2-5.

51. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR: Publication bias in clinical research. *Lancet* 1991;337:867-872.

52. Streiner DL, Norman GR, Cariney J: *Reliability (Chapter 8)*, in Streiner DL, Norman GR, Cariney J, eds: *Health Measurement Scales: A Practical Guide to Their Development and Use*, ed 5. Oxford, UK, Oxford University Press, 2015, pp 159-196.

53. Hung M, Stuart AR, Saltzman CL, Kubiak EN: Computerized adaptive testing using the PROMIS physical function item bank reduces test burden with less ceiling effects compared to the short musculoskeletal function assessment in orthopaedic trauma patients. *J Orthop Trauma* 2014;28:439-443.

54. Hibbard JH, Stockard J, Mahoney ER, Tusler M: Development of the Patient Activation Measure (PAM): Conceptualizing and measuring activation in patients and consumers. *Health Serv Res* 2004;39(4 Pt 1):1005-1026.

55. International Consortium for Health Outcomes Measurement (ICOMH). http://www.icohm.org. Accessed June 13, 2017.

56. PROMIS Health Organization, PROMIS Cooperative Group. PROMIS Instrument Development and Validation Scientific Standards Version 2.0. 2012 (Revised May 2013). http://www.nihpromis.org/Documents/ POMISStandards, Vers2.0_Final.pdf. Accessed July 20, 2017.

57. Revicki DA, Cell DF: Health status assessment for the twenty-first century: Item response theory, item banking and computer adaptive testing. *Qual Life Res* 1997;6:595-600.

58. 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.

59. Döring AC, Nota SP, Hageman MG, Ring DC: Measurement of upper extremity disability using the patient-reported outcomes measurement information system. *Hand Surg Am* 2014;39:1160-1165.

60. AO Trauma: http://www.aopatientoutcomescenter.com. Accessed July 20, 2017.

61. Hays RD, Spritzer KL, Amtmann D, Lai JS, Dewitt EM, Rthroock N: Upper-extremity and Mobility Subdomains from the Patient-Reported Outcomes Measurement Information System (PROMIS) adult physical functioning item bank. *Arch Phys Med Rehabil* 2013;94:2291-2296.

62. Cook KF, Roddey TS, Gartsman GM, Olson SL: Development and psychometric evaluation of the Flexilevel Scale of Shoulder Function. *Med Care* 2003;41:823-835.

63. van de Water AT, Davidson M, Shields N, Evans MC, Taylor NF: The shoulder function index (SFInX): A clinician-observed outcome measure for people with a proximal humeral fracture. *BMC Musculoskelet Disord* 2015;16:31.

64. Stuart AR, Higgins TF, Hung M, Weir CR, Kubiak EN, Rothberg DL: Reliability in measuring pre-injury physical function in orthopaedic trauma. *J Orthop Trauma* 2015;29:527-532.

65. Walton DM, MacDermid JC, Pullickal M, Rollack A, Veitch J: Development and initial validation of the Satisfaction and Recovery Index (SRI) for measurement of recovery from musculoskeletal trauma. *Open Orthop J* 2014;8:316-325.

66. Knutsen EJ, Paryavi E, Castillo RC, O’Toole RV: Is satisfaction among orthopaedic trauma patients predicted by depression and activation levels? *J Orthop Trauma* 2015;29:e183-e187.

67. Gruber JS, Hageman M, Neuhaus V, Mudgal CS, Jupiter JB, Ring D: Patient activation and disability in upper extremity illness. *J Hand Surg Am* 2014;39:1378-1383.

68. Porter ME, Larsson S: Standardizing patient outcomes measurement. *N Engl J Med* 2016;374:504-506.
### Appendix 1. PRISMA 2009 Checklist

| Section/Topic                  | No. | Checklist Item                                                                                                                                                                                                                                                                                                                                 | Reported on Page No. |
|-------------------------------|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|
| Title                         | 1   | Identify the report as a systematic review, meta-analysis, or both.                                                                                                                                                                                                                                                                           | 1                    |
| Abstract                      | 2   | Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.                                                                                  | 1                    |
| Introduction                  | 3   | Describe the rationale for the review in the context of what is already known.                                                                                                                                                                                                     | 3, 4                 |
| Objectives                    | 4   | Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).                                                                                                                                                                          | 4                    |
| Methods                       | 5   | Indicate if a review protocol exists, if and where it can be accessed (eg, Web address), and, if available, provide registration information including registration number.                                                                                                                                                                                                 | 5                    |
| Protocol and registration     | 6   | Specify study characteristics (eg, PICOS and length of follow-up) and report characteristics (eg, years considered, language, and publication status) used as criteria for eligibility, giving rationale.                                                                                                                                                                                  | 5                    |
| Eligibility criteria          | 7   | Describe all information sources (eg, databases with dates of coverage and contact with study authors to identify additional studies) in the search and date last searched.                                                                                                                                                                               | 5                    |
| Information sources           | 8   | Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.                                                                                                                                                                                                              | Supplemental Digital Content 1 (http://links.lww.com/JGB/A2)       |
| Search                        | 9   | State the process for selecting studies (ie, screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).                                                                                                                                                                                      | Figure 1             |
| Study selection               | 10  | Describe method of data extraction from reports (eg, piloted forms, independently, and in duplicate) and any processes for obtaining and confirming data from investigators.                                                                                                                                                                                    | 6                    |
| Data collection process       | 11  | List and define all variables for which data were sought (eg, PICOS and funding sources) and any assumptions and simplifications made.                                                                                                                                                                                                      | 6                    |
| Risk of bias in individual studies | 12 | Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.                                                                                                                                  | 6                    |
| Summary measures              | 13  | State the principal summary measures (eg, risk ratio and difference in mean).                                                                                                                                                                                                       | n/a                  |
| Synthesis of results          | 14  | Describe the methods of handling data and combining results of studies, if done, including measures of consistency (eg, $I^2$) for each meta-analysis.                                                                                                                                                                                      | 6                    |
| Risk of bias across studies   | 15  | Specify any assessment of risk of bias that may affect the cumulative evidence (eg, publication bias, selective reporting within studies).                                                                                                                                                                                                    | 6                    |

(continued)
| Section/Topic             | No. | Checklist Item                                                                                                                                                                                                 | Reported on Page No. |
|--------------------------|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|
| Additional analyses      | 16  | Describe methods of additional analyses (eg, sensitivity or subgroup analyses and meta-regression), if done, indicating which were prespecified.                                                                  | n/a                  |
| Results                  |     |                                                                                                                                                                                                                   |                      |
| Study selection          | 17  | Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.                                                   | 7                    |
| Study characteristics    | 18  | For each study, present characteristics for which data were extracted (eg, study size, PICOS, and follow-up period) and provide the citations.                                                                      | 7                    |
| Risk of bias within studies | 19  | Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).                                                                                                       | Table 2              |
| Results of individual studies | 20  | For all outcomes considered (benefits or harms), present, for each study: (1) simple summary data for each intervention group (2) effect estimates and confidence intervals, ideally with a forest plot.        | n/a                  |
| Synthesis of results     | 21  | Present the main results of the review. If meta-analyses are done, include for each, confidence intervals and measures of consistency.                                                                                | 7                    |
| Risk of bias across studies | 22  | Present results of any assessment of risk of bias across studies (see item 15).                                                                                                                                  | Table 2              |
| Additional analysis      | 23  | Give results of additional analyses, if done (eg, sensitivity or subgroup analyses, meta-regression [see item 16]).                                                                                                | n/a                  |
| Discussion               |     |                                                                                                                                                                                                                   |                      |
| Summary of evidence      | 24  | Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (eg, healthcare providers, users, and policy makers).                                      | 8                    |
| Limitations              | 25  | Discuss limitations at study and outcome level (eg, risk of bias) and at review-level (eg, incomplete retrieval of identified research and reporting bias).                                                        | 9                    |
| Conclusions              | 26  | Provide a general interpretation of the results in the context of other evidence and implications for future research.                                                                                             | 8, 13                |
| Funding                  |     |                                                                                                                                                                                                                   |                      |
| Funding                  | 27  | Describe sources of funding for the systematic review and other support (eg, supply of data); role of funders for the systematic review.                                                                           | n/a                  |

Reproduced with permission from Moher D, Liberati A, Tetzlaff J, Altman DG; The PRISMA Group: Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med* 2009 6(6):e1000097.