



Common Scales and Checklists in Sports Medicine Research

46

Alberto Grassi, Luca Macchiarola, Marco Casali,
Ilaria Cucurnia, and Stefano Zaffagnini

46.1 Introduction

Since the main purpose of clinical studies, especially randomized controlled trials (RCT), is to report or compare the effect of different treatments, the measurement methods of clinical outcomes are crucial. Therefore, during the early stage of study design, attention should be directed to choosing the appropriate outcomes and scales that evaluate a patient population. The calculation of sample size of a RCT is primarily based on the primary outcome being evaluated. When dichotomous outcomes of rare events such as failures or complications are used, extremely large sample sizes are often required. This requirement may discourage the realization of the study or require

an enormous amount of resources to reach adequate enrollment. Conversely, if continuous measures such as patient- or clinician-reported scales are chosen, a power analysis based on means and standard deviations usually provides more feasible sample size.

However, care should be used during the power analysis to ensure that it is based on a clinical score which is able to detect a real difference between the different treatments. In fact, building a clinical study or RCT on outcomes which are not completely appropriate to the study purpose, patient population, and treatment administered could compromise the utility and subsequent impact of the results. Therefore, the researcher should be very familiar with the main features of clinical scores and should also know the main characteristics of each scale in relation to the pathology or treatment that is being investigated.

Another important aspect in outcome selection is the global assessment of the patient. Traditionally, clinical outcomes in orthopedics consisted of measuring impairments such as range of motion, joint stability, strength, pain, and joint function. At times, surgeons are marginally interested in patient's global disability and mental status; however, the patient's perception of changes in health status is the most important indicator of the success of a treatment. Therefore, there are two possibilities of measuring health-related quality of life in orthopedic and sports medicine conditions. The "generic measures"

A. Grassi
IIa Clinica Ortopedica e Traumatologica, IRCCS
Istituto Ortopedico Rizzoli, Bologna, Italy

Dipartimento di Scienze Biomediche e Neuromotorie
(DIBINEM), Università di Bologna, Bologna, Italy

SIGASCOT Arthroscopy Committee, Florence, Italy
e-mail: alberto.grassi@ior.it

L. Macchiarola (✉) · M. Casali · I. Cucurnia
IIa Clinica Ortopedica e Traumatologica, IRCCS
Istituto Ortopedico Rizzoli, Bologna, Italy

S. Zaffagnini
IIa Clinica Ortopedica e Traumatologica, IRCCS
Istituto Ortopedico Rizzoli, Bologna, Italy

Dipartimento di Scienze Biomediche e Neuromotorie
(DIBINEM), Università di Bologna, Bologna, Italy
e-mail: stefano.zaffagnini@unibo.it

pertain the overall health of the patient, including physical, mental, and social well-being, and offer the advantage of being able to use them to compare different diseases, severity, and interventions. However, since they represent a generic measure, their ability to detect small but important changes could be limited. On the other hand, the “disease-specific measures,” which pertain to a specific disorder treated in a patient, measure the physical, mental, and social aspects of health affected by the specific disorder. Therefore, they are able to detect small but important changes but have a limited value in comparison of health status across different diseases. For the aforementioned reasons, a complete picture of treatment effect on a patient could be provided only with the assessment through a “disease-specific measure” in combination with a “generic measure.”

Fact Box 46.1

Too often surgeons are poorly interested in patient’s global health self-perception and mental status; however, the patient’s perception of changes in health status and disability is the main indicator of the success of a treatment.

46.2 General Scale Characteristics

The main characteristics and features of clinical scales that should be known to choose the appropriate outcome measure are the following [54].

Construct validity: is defined as the ability of an instrument to measure what it is supposed to measure. It depends on how the items that make up the scale include all relevant aspects of the pathology or disability that is measured. The convergent validity indicates how the score could correlate with other scores that measure the same construct. Meanwhile, predictive validity indicates whether the score could predict a patient’s score on a measure of some related construct.

Repeatability (Test/retest reliability): is defined as the agreement between the observations on the same patients on two or more occa-

sions separated by a time interval under stable health conditions. It is considered when the raters are not involved or the raters’ effect is negligible. It could be assessed with the intraclass correlation (ICC) or the Cohen’s *K* statistics.

Intra-rater reliability: is defined as the agreement between two or more repeated score evaluation performed by a single rater. Also in this case, intraclass correlation (ICC) or the Cohen’s *K* statistics could be used.

Inter-rater reliability: is defined as the agreement between the scores obtained from two or more raters’ assessment. It measures how much consensus or heterogeneity there is in the rating given by judges. Similarly, intraclass correlation (ICC) or the Cohen’s *K* statistics is employed.

Internal consistency: is defined as the correlations between different items on the same test and measures whether several items that propose to measure the same general construct produce similar scores. It is assessed using the Cronbach’s alpha statistics.

Responsiveness to change: is defined as the ability of an instrument to detect clinically important changes between the patient’s pre-intervention and post-intervention state, assuming all other factors remain constant.

Minimal detectable change (MDC): is defined as the minimal change that falls outside the measurement error in the score of an instrument used to measure a symptom.

Minimal clinically important difference (MCID): is defined as the minimal change in the score that is meaningful for patients or that is required for the patient to feel a difference in the variable that is measured.

Standard error of measurement (SEM): It measures the range within which a score would likely fall in the case of re-measurement.

Standardized response mean (SRM): It measures the responsiveness to change and is defined as the mean change in score divided by the standard deviation of the change scores.

Floor effect: Floor effects occur when a measure’s lowest score is unable to assess a patient’s level of ability. The test is considered poor if the floor effect is >20%.

Ceiling effect: Ceiling effects occur when a measure's highest score is unable to assess a patient's level of ability. This might be particularly common for measures used over multiple occasions. The test is considered poor if the ceiling effect is >20%.

46.3 Measures of Shoulder Function

There are many instruments that measure symptom and function of the shoulder and some that evaluate both the glenohumeral joint and the whole upper limb. The most widespread and best tested is the disabilities of the arm, shoulder, and hand questionnaire (DASH). Also, the shoulder pain and disability index (SPADI), the Constant-Murley score (CMS), and the American shoulder and elbow surgeons (ASES) questionnaire, which are more specific for shoulder pathologies, are extensively employed. The simple shoulder test (SST), the shoulder disability questionnaire (SDQ), the Oxford shoulder score, and the West Ontario shoulder instability index (WOSI) complete the panorama of most common tools.

The basic psychometric characteristics, strengths, and weaknesses of the most common scales for shoulder function are described (Table 46.1).

46.3.1 Disabilities of the Arm, Shoulder, and Hand Questionnaire (DASH)

The DASH is a patient-completed scale which includes 30 items regarding symptoms, pain, physical function, and social function [58]. The 11-item QuickDASH short version is also available [5]. The DASH is the best tool for the comprehensive assessment of upper extremity conditions, since it is easy to apply, analyze, and interpret; moreover it is good for research purposes in various upper extremity conditions and has a good correlation with SPADI, HAQ, CMS, ASES, and EQ-5D with Pearson's or Spearman's

test. It is particularly useful when polyarticular conditions should be evaluated or symptoms and function of the entire upper extremity are investigated. It is also useful in all elbow and hand conditions. However, the DASH is region specific and not joint specific; therefore specificity and responsiveness are lower than those of unique shoulder-specific tools [3].

Conditions: Any or multiple disorders of upper extremity, in particular painful conditions including: rheumatoid arthritis, multiple sclerosis, adhesive capsulitis, shoulder impingement and tendinitis, proximal humerus fracture, distal radius fractures, hand osteoarthritis or fractures, arthroscopic acromioplasty.

46.3.2 Shoulder Pain and Disability Index (SPADI)

The SPADI is a patient-completed scale that includes 13 items regarding symptoms and pain, scored on a VAS/NRS scale [94]. It is one of the most representative shoulder instruments and has been tested in numerous settings; moreover, it is easy to administer, understand, and complete. It has a good correlation with the DASH, ASES, and CMS. One possible weakness in construct validity could be that only one item assesses overhead work [3].

Conditions: Any disorder of the shoulder joint, particularly adhesive capsulitis, rotator cuff pathologies.

46.3.3 American Shoulder and Elbow Surgeons Society Shoulder Assessment Form (ASES)

The ASES is a patient self-evaluation scale of 11 items evaluation pain and function, which is integrated with a clinician-dependent part [92]. It has good reliability, construct validity, and responsiveness. However, it uses different type of scales (binary, Likert, VAS), and the clinician part could be time-consuming. It has been developed to be applied to all shoulder patients regardless of the diagnosis, since it evaluates also activities of

Table 46.1 Measures of shoulder function

| Questionnaire | Items | Options | Administration | Range | Cutoffs | Time compilation (min) | Time calculation (min) | Internal consistency | Test-retest | SRM | MDC | MCID |
|--------------------|-------|---------------|---------------------|-------------------------|---------|------------------------|------------------------|----------------------|-------------|-----------|-----------|-----------|
| DASH ^a | 30 | Likert (5) | Patient | 0 (best) to 100 (worst) | Yes | 4 | 10 | 0.92–0.98 | 0.93–0.98 | 0.43–1.2 | 7.9–14.8 | 10.2 |
| SPADJ ^a | 13 | VAS/NRS | Patient | 0 (worst) to 10 (best) | No | 5 | 2 | 0.86–0.96 | 0.84–0.95 | 1.23–1.81 | 13.2–21.5 | 13.2–23.1 |
| ASES ^a | 11 | Mix | Patient | 0 (worst) to 100 (best) | No | 3 | >8 | 0.61–0.96 | 0.84–0.9 | 1.42–1.81 | 11.2 | 6.4–16.9 |
| CMS ^a | 8 | Likert (3–10) | Patient + clinician | 0 (worst) to 100 (best) | Yes | 5 | NA | 0.60 | 0.80–0.96 | 0.59–2.09 | NR | NR |
| SST ^b | 12 | Dichotomous | Patient | 0 (worst) to 12 (best) | No | 2 | 5 | 0.85 | 0.97–0.99 | 0.63–1.94 | NR | 2.05–2.33 |
| OSS ^a | 12 | Likert (5) | Patient | 12 (best) to 60 (worst) | No | 3 | 5 | 0.94 | 0.98 | 1.10–1.14 | NR | NR |
| UCLA ^b | 5 | Likert (2–6) | Patient + clinician | 0 (worst) to 35 (best) | Yes | 5 | 5 | 0.93–0.95 | NR | 0.15–0.90 | NR | NR |
| WOSI ^a | 21 | VAS/NRS | Patient | 0 (best) to 100 (worst) | No | 3 | 6 | 0.88–0.96 | 0.87–0.98 | 0.93–1.40 | NR | NR |
| WOOS ^c | 19 | VAS | Patient | 0 (worst) to 100 (best) | No | 10 | 5 | NR | 0.96 | 1.91 | NR | NR |
| WORC ^d | 20 | VAS | Patient | 0 (worst) to 100 (best) | No | 10 | 5 | NR | 0.96 | NR | NR | 11.7 |
| OSIQ ^e | 12 | Likert (5) | Patient | 0 (worst) to 48 (best) | Yes | 5 | 5 | 0.88 | 0.87 | NR | 9 | NR |

Note: DASH (disabilities of the arm, shoulder, and hand questionnaire), SPADI (shoulder pain and disability index), ASES (American shoulder and elbow surgeons society shoulder assessment form), CMS (Constant-Murley score), SST (simple shoulder test), OSS (Oxford shoulder score), UCLA (University of California at Los Angeles shoulder score), WOSI (Western Ontario shoulder instability index), WOOS (Western Ontario osteoarthritis of the shoulder index), WORC (Western Ontario rotator cuff index), OSIQ (Oxford shoulder instability questionnaire)

^aAngst F, Schwytzer HK, Aeschlimann A, Simmen BR, Goldhahn J. Measures of adult shoulder function. *Arthritis Care Res* (Hoboken). 2011 Nov;63 Suppl 11:S174–88

^bAmstutz HC, Sew Hoy AL, Clarke IC. UCLA anatomic total shoulder arthroplasty. *Clin Orthop Relat Res*. 1981 Mar-Apr;(155):7–20

^cLo IK, Griffin S, Kirkley A. The development of a disease-specific quality of life measurement tool for osteoarthritis of the shoulder: The Western Ontario Osteoarthritis of the Shoulder (WOOS) index. *Osteoarthritis Cartilage*. 2001 Nov;9(8):771–8

^dKirkley A, Griffin S, Dainty K. Scoring systems for the functional assessment of the shoulder. *Arthroscopy*. 2003 Dec;19(10):1109–20

^evan der Linde JA, van Kampen DA, van Beers LW, van Deursen DF, Terwee CB, Willems WJ. The Oxford Shoulder Instability Score; validation in Dutch and first-time assessment of its smallest detectable change. *J Orthop Surg Res*. 2015 Sep 17;10:146

daily living. It has a good correlation with both SPADI and DASH questionnaires [3].

Conditions: Any disorder of the shoulder joint, particularly rotator cuff disease, shoulder impingement, shoulder arthritis, calcific tendonitis.

46.3.4 Constant-Murley Score (CMS)

The CMS is both patient- and clinician-reported score which includes eight items regarding pain, ADLs, mobility, and strength. It is a method to record individual parameters, providing an overall clinical functional assessment, irrespective of diagnosis or radiographic abnormalities [22, 23]. Based in the difference with the abnormal side, the indexed shoulder could be graded as excellent (<11), good (11–20), fair (21–30), or poor (<30). Despite the CMS is highly accepted throughout the clinical community, there are several limitations to its use due to the low inter-tester reliability, non-standardized measurement of strength, and only few items evaluating pain and ADL. It is useful for measurement protocols but does not provide an adequate self-assessment of patient pain and function. It has a good correlation with ASES, DASH, and SPADI [3].

Conditions: Mainly rotator cuff-related disorders, impingement, degenerative or inflammatory pathologies, instability, osteoarthritis.

46.3.5 Simple Shoulder Test (SST)

The SST is a patient-reported score which includes 12 dichotomous (yes/no) items regarding pain, strength, and range of motion [71]. It assesses the functional disability of the shoulder in a very simple and short manner; however, due to the binary response option, its use as a comprehensive measure of outcomes could be questioned. It has a good correlation with the SPADI, ASES, DASH, and CMS scores [3].

Conditions: General shoulder injuries and rotator cuff pathology.

46.3.6 Oxford Shoulder Score (OSS)

The OSS is a patient-reported scale of 12 items evaluating pain and daily function [32, 35]. It provides a self-assessment of shoulder pain and function. It is short and easy to complete but not frequently used in the current literature. Correlation with SPADI, DASH, and CMS is good [3].

Conditions: Degenerative and inflammatory shoulder conditions, subacromial impingement, rotator cuff, osteoarthritis, and proximal humerus fractures.

46.3.7 UCLA Shoulder Score

The UCLA (University of California at Los Angeles) shoulder score is both a five-item patient- and clinician-reported scale which evaluates pain, function, ROM, strength, and patient's satisfaction [2]. Despite being one of the earliest available shoulder outcome measures, it has not formally been validated. It is simple and fast but requires physician manual evaluation; for this reason, it could result in a poor validity or responsiveness, which does not make it ideal for research setting. The UCLA has a good correlation with the DASH, SPADI, and SF-36 and could be dichotomized as good/excellent (>27) or fair/poor (<27) [61].

Conditions: Common shoulder pathologies.

46.3.8 Western Ontario Shoulder Instability Index (WOSI)

The WOSI is a 21-item patient-reported scale that evaluates physical symptoms, pain, sport, work, lifestyle, and emotions related to shoulder instability [62, 63]. It has been developed to assess disease-specific quality of life patients with symptomatic shoulder instability. It has the advantage of being specific for this condition, but due to lack of testing data, caution is necessary at individual patient level. It has a good correlation with the VAS for function and the DASH score [3].

Conditions: Shoulder instability.

46.3.9 Western Ontario Osteoarthritis of the Shoulder Index (WOOS)

The WOOS is a 19-item patient-reported questionnaire that evaluates the area of pain, physical symptoms, sports and work, lifestyle function, and emotional function [72]. Its form as 100-mm VAS makes it an easy, fast, and reliable questionnaire; however, it is specific for degenerative pathologies, especially osteoarthritis. Its multiple domains regarding both function and psychologic aspects make the WOOS a versatile and complete scale. In fact, it contains many items rarely investigated by other shoulder questionnaires. It has a moderate correlation with the Constant-Murley and UCLA scores [61].

Conditions: Osteoarthritis of the shoulder.

46.3.10 Western Ontario Rotator Cuff Index (WORC)

The WORC is a 20-item patient-reported score that evaluates symptoms, sport, work, emotion, and social function [60]. It is easy and rapid to administer, since it is composed of 100-mm VAS items, but it is disease-specific since it has been developed to evaluate rotator cuff-related quality of life.

It has a good correlation with the ASES and UCLA scores [61].

Conditions: Rotator cuff pathology treated surgically and conservatively.

46.3.11 Oxford Shoulder Instability Questionnaire (OSIQ)

The OSIQ is a 12-item patient-reported questionnaire that explores the impact of shoulder instability on work, sport and social life, its psychological repercussion, the quality of life, and the pain [33, 35]. It is specifically designed for glenohumeral dislocation and shoulder instability. However, it has a good correlation with both DASH and WOSI. Based on the obtained value, function could be graded as

excellent (40–48), good (30–39), fair (20–29), or poor (0–19) [108].

Conditions: Surgery or physiotherapy for shoulder instability. Shoulder instability.

46.4 Measures of Elbow, Wrist, and Hand Function

Elbow, wrist, and hand function represent a complex dimension to evaluate. Especially for elbow, physical examination and objective evaluation of ROM and stiffness are important characteristics to assess the joint function, patient's satisfaction, and normal or pathologic conditions. Therefore, clinical scores often require a clinician-reported items that increase the precision of the evaluation, but on the other side, they reduce the reliability and make them time-consuming as well.

The basic psychometric characteristics, strength, and weakness of the most common scales for elbow, wrist, and hand function are described (Table 46.2).

46.4.1 Mayo Elbow Performance Score (MEPS)

The MEPS is both a patient- and clinician-reported score. It includes four Likert-scale items evaluating mostly pain and motion, stability, and function [82]. It is correlated with other elbow measures for raw scores rather than categorical ranks and requires clinician objective evaluation of the patient, which could lengthen its application [16].

Conditions: General elbow disorders, rheumatoid arthritis, synovectomy.

46.4.2 Oxford Elbow Score (OES)

The OES is a 12-item patient-reported score. It includes 12 Likert-scale items evaluating elbow function, pain, and the psychological aspects [30]. It is easy and simple to administer to patients; however, it lacks objective evaluation of clinical outcomes. It has a good correlation with DASH, Mayo elbow score, and SF-36 [73].

Conditions: General elbow disorders.

Table 46.2 Measures of elbow, wrist, hand function

| Questionnaire | Items | Options | Administration | Range | Cutoffs | Time compilation (min) | Time calculation (min) | Internal consistency | Test-retest | SRM | MDC | MCID |
|--------------------|-------|--------------|---------------------|-------------------------|---------|------------------------|------------------------|----------------------|-------------|-----------|------|------|
| MEPS ^a | 4 | Likert (3–4) | Patient + clinician | 0 (worst) to 100 (best) | No | 5 | 3 | NR | 0.89 | NR | 11.3 | 15 |
| OES ^b | 12 | Likert (5) | Patient | 0 (worst) to 100 (best) | No | 3 | 3 | 0.8–0.90 | 0.87 | 0.46–0.69 | 27.6 | NR |
| ASES ^c | 57 | Mix | Patient + clinician | 0 (worst) to 100 (best) | No | 15 | 10 | 0.68–0.82 | 0.95 | NR | NR | NR |
| PRTEE ^d | 10 | NRS | Patient | 0 (best) to 100 (worst) | No | 3 | 3 | NR | 0.87 | 2.01 | NR | NR |
| MWS ^e | 4 | Likert (3–4) | Patient + clinician | 0 (worst) to 100 (best) | No | 5 | 5 | NR | NR | NR | NR | NR |
| MHQ ^f | 37 | Likert (5) | Patient | 0 (worst) to 100 (best) | No | 15 | 20 | 0.75–0.94 | 0.95 | 0.47–1.61 | NR | 3–13 |
| FIHOA ^f | 10 | Likert (4) | Patient | 0 (best) to 30 (worst) | No | 3 | 3 | 0.85–0.90 | 0.95 | 0.58–0.87 | NR | NR |

Note: MEPS (Mayo elbow performance score), OES (Oxford elbow score), ASES (American shoulder and elbow surgeons society shoulder assessment form), PRTEE (patient-rated tennis elbow evaluation), MWS (Mayo wrist score), MHQ (Michigan hand outcome questionnaire), FIHOA (functional index for hand osteoarthritis)

^aCelik D. Psychometric properties of the Mayo Elbow Performance Score. *Rheumatol Int.* 2015 Jun;35(6):1015–20

^bDawson J, Doll H, Boller I, Fitzpatrick R, Little C, Rees J, Jenkinson C, Carr AJ. The development and validation of a patient-reported questionnaire to assess outcomes of elbow surgery. *J Bone Joint Surg Br.* 2008 Apr;90(4):466–73

^cMacDermid JC. Outcome evaluation in patients with elbow pathology: issues in instrument development and evaluation. *J Hand Ther.* 2001 Apr-Jun;14(2):105–14

^dRompe JD, Overend TJ, MacDermid JC. Validation of the Patient-rated Tennis Elbow Evaluation Questionnaire. *J Hand Ther.* 2007 Jan-Mar;20(1):3–10

^eAmadio PC, Berquist TH, Smith DK, Ilstrup DM, Cooney WP 3rd, Linscheid RL. Scaphoid malunion. *J Hand Surg Am.* 1989 Jul;14(4):679–87

^fPoole JL. Measures of hand function. *Arthritis Care Res (Hoboken).* 2011 Nov;63 Suppl

46.4.3 American Shoulder and Elbow Surgeons Society Shoulder Assessment Form (ASES)

The ASES elbow outcome score is both a patient- and clinician-reported questionnaire that evaluates elbow pain, function, and satisfaction through 19 items and motion, stability, strength, and physical findings through 38 items [92]. This score represents a complete evaluation but requires substantial time to be complete, and pain has the highest influence in the overall score [80].

Conditions: General elbow disorders.

46.4.4 Patient-Rated Tennis Elbow Evaluation (PRTEE)

The PRTEE is a 15-item patient-reported score that evaluates forearm pain and disability in patients with lateral epicondylitis. It presents two subscales: pain and function. It is easy to complete and fast to be administered and had good correlation with NRS for pain during wrist extension and with the DASH; however, it is specific one condition [75, 97].

Conditions: Lateral epicondylitis of the elbow.

46.4.5 Mayo Wrist Score (MWS)

The MWS is both a patient- and clinician-reported score, which includes five Likert-scale items evaluating pain, function, ROM, and grip strength. It's basic but involves objective evaluation of wrist mobility and strength which could limit its usability [25]. Moreover, its reliability and consistency characteristics have not been deeply investigated. It could be graded as excellent (90–100), good (80–90), satisfactory (60–80), and poor (<60) [28].

Conditions: Originally developed for scaphoid nonunion; could be used for wrist fractures and arthritis.

46.4.6 Michigan Hand Outcome Questionnaire (MHQ)

The MHQ is a 37-item patient-reported scale that evaluates hand function, appearance, pain, and

satisfaction [19]. It appropriately measures hand function in various conditions; however, its application could be time-consuming [28].

Conditions: Hand and wrist injuries, including osteoarthritis.

46.4.7 Functional Index for Hand Osteoarthritis (FIHOA)

The FIHOA is a patient-reported scale composed of ten items including questions about using keys, cutting, lifting, buttoning, and writing, aimed to measure hand function in patients with hand osteoarthritis. It has a good correlation with the MHOQ [36].

Conditions: Hand osteoarthritis.

46.5 Measures of Hip Function

The assessment of outcomes in hip surgery is focused on patient satisfaction and the quality of life achieved, level of pain, range of motion (ROM), comorbidities, and the use of walking aids. A variety of quality of life evaluation tools have been developed that differ in their measurement techniques and in the number of domains they assess. These scores are useful not only for the normal clinical evaluation in old patients and in hip congenital disease but also to assess the outcomes after joint-preserving surgery. The ideal hip outcome measure should be one that is specific for the hip joint, possesses a generic component, and is clear and concise. Previous outcome tools were modifications of preexisting tools that evaluate chronic conditions such as osteoarthritis. Outcome measures most frequently used in clinical practice are the Harris hip score, the hip disability and osteoarthritis outcome score, the Oxford hip score, and the Lequesne index of severity for osteoarthritis of the hip. More specific scores for sport-related hip injuries were designed in the last years such as non-arthritic hip score and international hip outcome tool-33.

The basic psychometric characteristics, strength, and weakness of the most common scales for hip function are described (Table 46.3).

Table 46.3 Measures of hip function

| Questionnaire | Items | Options | Administration | Range | Cutoffs | Time compilation (min) | Time calculation (min) | Internal consistency | Test-retest | SRM | MDC | MCID |
|-----------------------|-------|---------------|---------------------|-------------------------|---------|------------------------|------------------------|----------------------|-------------|-----------|-----------|------|
| HHS ^a | 10 | Likert (2–15) | Clinician | 0 (worst) to 100 (best) | Yes | 5 | 10 | NR | 0.93–0.98 | 2.52–2.73 | NR | NR |
| HOOS ^b | 40 | Likert (5) | Patient | 0 (worst) to 100 (best) | No | 10–15 | 2–3 | 0.82–0.98 | 0.75–0.97 | 1.29–3.24 | 9.6–16.2 | NR |
| OHS ^c | 12 | Likert (5) | Patient | 0 (worst) to 48 (best) | Yes | 2–8 | 5 | 0.84–0.93 | 0.84–0.93 | 1.12 | 6.11 | 3–5 |
| LISOH ^a | 11 | Likert (2–7) | Patient + clinician | 0 (best) to 24 (worst) | Yes | 2–5 | 5 | 0.83–0.84 | 0.94 | NR | NR | NR |
| NAHS ^c | 20 | Likert (5) | Patient | 0 (worst) to 100 (best) | No | 8–10 | 5 | 0.69–0.92 | 0.87–0.95 | NR | 10.4–12.4 | NR |
| iHOT33 ^{d,e} | 33 | VAS | Patient | 0 (worst) to 100 (best) | No | 15 | 6 | 0.96–0.99 | 0.87–0.96 | 1.7 | 16.0 | 6.1 |
| HAGOS ^f | 37 | Likert (5) | Patient | 0 (worst) to 100 (best) | No | 10 | 10 | 0.37–0.73 | 0.82–0.92 | NR | 02-mag | NR |

Note: HHS (Harris hip score), HOOS (hip disability and osteoarthritis outcome score), OHS (Oxford hip score), LISOH (Lequesne in severity for osteoarthritis of the hip), NAHS (non-arthritic hip score), iHOT33 (international hip outcome tool-33), HAGOS (the Copenhagen hip and groin outcome score)

^aNilsdotter A, Bremander A. Measures of Hip Function and Symptoms. *Arthritis Care & Research* Vol. 63, No. S11, November 2011, pp. S200–S208

^bMartinelli N, Longo UG, et al. Cross-cultural adaptation and validation with reliability, validity, and responsiveness of the Italian version of the Oxford Hip Score in patients with hip osteoarthritis. *Qual Life Res* (2011) 20:923–929

^cRamisetty N, Kwon Y, and Mohrtdi N. Patient-reported outcome measures for hip preservation surgery. A systematic review of the literature. *Journal of Hip Preservation Surgery* Vol. 2, No. 1, pp. 15–27

^dMohrtdi NG, Griffin DR, Pedersen ME et al. The development and validation of a self-administered quality-of-life outcome measure for young, active patients with symptomatic hip disease: the international hip outcome tool (iHOT-33). *Arthroscopy* 2012; 28: 595; 605

^eKemp JL, Collins NJ, Roos EM et al. Psychometric properties of patient-reported outcome measures for hip arthroscopic surgery. *Am J Sports Med* 2013; 41: 2065–73

^fThorborg K, Holmich P, Christensen A, Petersen F, Roos EM. The Copenhagen Hip and Groin Outcome Score (HAGOS): development and validation according to the COSMIN checklist. *Br J Sports Med*. 2011;45:478–491

46.5.1 Harris Hip Score (HHS)

The HHS is a clinician-based outcome score which includes ten items that evaluates pain, function, absence of deformity, and range of motion [44]. There are two versions of the score: the original one, published in 1969, and the modified HHS (MHHS). The latter only includes pain and function components and has been widely used to evaluate outcomes in hip arthroscopy surgery. The HHS is widely used throughout the world for evaluating outcome after THR, and it has also been proven appropriate to measure outcome after surgical interventions for femoral neck fractures. It seems to be useful for short-time follow-up; moreover there are unacceptable ceiling effects that severely limit its validity. The HHS has been used in many different countries (Sweden, the Netherlands, Denmark, etc.), but there are no validated versions in other languages available. It has a good correlation with WOMAC, NHP, NAHS, and SF-36 for pain and function domain. Based on the obtained score, it could be graded as excellent (90–100), good (80–90), fair (70–80), or poor (<70) [86].

Conditions: Femoral neck fractures, osteoarthritis of hip, impingement.

46.5.2 Hip Disability and Osteoarthritis Outcome Score (HOOS)

The HOOS is a patient-reported score composed of 40 items that evaluates pain, other symptoms, function in activities of daily living, function in sport and recreation, and hip-related quality of life. It has been validated in two slightly different versions, LK1.1 and LK2.0 [65, 85]. In 2008, a five-item measure of physical function, the HOOS-PS, was published derived from the HOOS questionnaire by item response theory to elicit patients' opinions about difficulties experienced due to hip problems. The HOOS has been used in subjects with hip disability with or without hip osteoarthritis and in patients with hip osteoarthritis pre- and postoperative total hip

replacement. The HOOS is an extension of the WOMAC and is suggested to be valuable for younger and more active people due to the added subscales. It is suitable for use in research as a disease-specific questionnaire. It has a good correlation with Oxford hip score, the Lequesne index, and the VAS for pain. Based on its score, it could be graded as excellent (>41), good (34–41), fair (27–33), or poor (<27) [86].

Conditions: Osteoarthritis, general hip disorders.

46.5.3 Oxford Hip Score (OHS)

The OHS is a 12-item patient-reported outcome score regarding pain and function of the hip in relation to daily activities such as walking, dressing, and sleeping. It is designed for the assessment of joint replacement and has been used in several countries in large registry studies [31, 83]. The OHS was developed to supplement other generic outcome measures in systematic studies of hip replacement surgery with long-time follow-up. It has also been validated and used in revision hip replacement. Due to its shortness, the OHS questionnaire is feasible for surveys by mail, and it yields a high response rate and is therefore preferred for larger studies. High correlation was found between OHS and the HHS in THR patients [86].

Conditions: Osteoarthritis of hip.

46.5.4 Lequesne in Severity for Osteoarthritis of the Hip (LISOH)

The LISOH is an interview-based or reported score which includes 11 items evaluating pain, maximum distance walked, and activities of daily living [68–70]. The LISOH is available currently in several versions: interview based, self-administered, and in modified versions due to changed scoring and wording. The LISOH was developed to evaluate the severity of hip osteoarthritis in drug trials and the long-term treatment effects for hip OA and as help in decision-making regarding

the need for hip replacement. It has limited construct validity; also the convergent validity of the questionnaire has been questioned. Recommendations are to only use the LISOH for group comparisons. Based on its score, the handicap derived from hip osteoarthritis could be graded as extremely severe (>14), very severe (11–13), severe (10–8), moderate (5–7), mild (1–4), or none (0) [86].

Conditions: Osteoarthritis, the effectiveness of pharmacologic interventions, and to help with indications for surgery like THA.

46.5.5 Non-arthritic Hip Score (NAHS)

The non-arthritic hip score (NAHS) consists of 20 items distributed in four domains of pain, mechanical symptoms, functional symptoms, and activity level. It was developed for young active patients with higher demands and expectations. This is a patient-based, self-administered questionnaire that was developed as a modification of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). Input from patients, surgeons, physical therapists, and epidemiologists was used in creating NAHS scoring system. The NAHS has satisfactory internal consistency in each of its four domains. But there is no further evidence about internal consistency from head-to-head comparison studies with other outcome measures. Hence, the summation score for internal consistency for NAHS is good. The summation score for test-retest reliability is excellent. Construct validity is satisfactory between the NAHS and the Harris hip score (HHS) and short form (SF)-12, respectively [18].

Conditions: All non-arthritic hip conditions.

46.5.6 International Hip Outcome Tool-33 (iHOT-33)

These 33 questions were formulated into a self-administered questionnaire using a visual analog scale response format from 0 to 100 (worst–best

outcome) [81]. The iHOT-33 was developed with the cooperation of the multicenter arthroscopy of the hip outcomes research network (MAHORN). It has a short version: the iHOT12 that includes 12 items instead of the original 33, designed to be more easily used in clinical settings and validated and tested for reliability. The appropriate population for this tools includes patients aged between 18 and 60 years who have a Tegner activity level of 4 or higher, meaning that they are engaged in recreational physical activities at least once a week or have an occupation involving moderately heavy labor. There were no floor or ceiling effects noted for iHOT-33 in their original paper. In the end the construct validity was demonstrated with a correlation of 0.81 to the NAHS [57].

Conditions: Femoroacetabular impingement, hip arthroscopic surgery for intra-articular hip lesion.

46.5.7 The Copenhagen Hip and Groin Outcome Score (HAGOS)

The HAGOS is a patient-reported outcome questionnaire; it consists of 37 items distributed in six subscales of pain, symptoms, physical function in activities of daily living, physical function in sports and recreation, participation in physical activities, and hip- and/or groin-related quality of life [107].

The Copenhagen hip and groin outcome score was developed in 2011, and this was the first outcome measure developed with the COSMIN checklist guidelines. The goal of this instrument is to evaluate hip and/or groin disability related to impairment (body structure and function), activity (activity limitations), and participation (participation restrictions) according to the international classification of functioning, disability, and health (ICF) in young to middle-aged physically active patients with hip and/or groin pain. The HAGOS has excellent test-retest reliability properties; this was evident from ICC ranging from 0.82 to 0.92 for all its subscales from their original paper, and it has also excellent

internal consistency properties and good content validity. Floor or ceiling effects were noted in some subscales of HAGOS as described in their original paper, while there were no floor effects for HAGOS ceiling effects that were noted in HAGOS ADL (32%) and physical activity (28%) subscales between 12 and 24 months after surgery. In the end construct validity is satisfactory between the HAGOS and the SF-36 [57, 91].

Conditions: Young to middle-aged patients with long-standing hip and/or groin pain.

46.6 Measures of Knee Function

The knee is one of the most investigated joints in Orthopaedics and Sports Medicine; therefore many outcome measures exist for clinical or research use. The most relevant are those evaluating pain, function, quality of life, and activity. However, the clinician-reported scales that have been described allow to register objective characteristics of the joint such as deformity, ROM, and stability.

The basic psychometric characteristics, strength, and weakness of the most common scales for knee function are described (Table 46.4).

46.6.1 International Knee Documentation Committee Subjective Score (Subjective IKDC)

The subjective IKDC form is an 18-item patient-reported score that examines knee symptoms, sport participation, and daily activities [53]. It was developed in 1994 and revised to its current form in 2001. Its strength is the comprehensive assessment of the patient status and, above all, responsiveness to change following surgical interventions. Its limited administrative and respondent burden makes it ideal in both clinical and research settings. Moreover, it has a good correlation with the Cincinnati knee rating system, VAS for pain, WOMAC, Lysholm, and SF-36. On the other hand, it lacks in the psychometric testing, which makes it suboptimal for the evaluation of osteoarthritic patients [21].

Conditions: Knee ligament injury and surgery, meniscal tears, cartilage lesions, knee dislocation.

46.6.2 International Knee Documentation Committee Objective Score (Objective IKDC)

The objective IKDC form is a clinician-reported score that evaluates all the aspects of knee findings. Twenty-five items, grouped in seven subgroups evaluating effusion, passive motion deficit, ligament examination, crepitus, harvest site pathology, X-ray findings, and functional evaluation with one-leg hop test, compose it. Every item is rated in a four-grade Likert scale from A to D or normal to severely abnormal, with respect to the contralateral healthy knee. The overall score is determined by the lowest value of considering only the first three subgroups (swelling, passive ROM, and ligament stability). However, all the items should be compiled even if they did not contribute to the overall score. This form is frequently used when evaluating ligamentous surgery and allows the comparison of different groups of treatment in a reliable manner. However, to increase its precision, it requires instrument-assisted evaluation of knee stability. Moreover, in the case of bilateral pathology or contralateral previous injury, the score could not be used since it implies the comparison to a healthy contralateral limb. Usually, the grade C and D are considered as failure of the treatment [1].

Conditions: Especially knee ligament injury and surgery, but also other knee conditions investigated by subjective IKDC form such as meniscal tears, cartilage lesions, knee dislocation.

46.6.3 Knee Injury and Osteoarthritis Outcome (KOOS)

The KOOS is a 42-item patient-reported scale, which includes five domains, each one scored separately: pain, symptoms, activity of daily life (ADL), sports and recreational activities, and

Table 46.4 Measures of knee function

| Questionnaire | Items | Options | Administration | Range | Cutoffs | Time compilation (min) | Time calculation (min) | Internal consistency | Test-retest | SRM | MDC | MCID |
|------------------------------|-------|---------------|---------------------|---|---------|------------------------|------------------------|----------------------|-------------|-----------|-----------|-------|
| Subjective IKDC ^a | 18 | Likert (5–10) | Patient | 0 (worst) to 100 (best) | No | 10 | 5 | 0.92–0.97 | 0.87–0.89 | 0.94–1.5 | 6.7 | 11.5 |
| Objective IKDC ^b | 25 | Likert (4) | Clinician | A (best) to D (worst) | Yes | 15 | 3 | NR | NR | NR | NR | NR |
| KOOS ^a | 25 | Likert (5) | Patient | 0 (worst) to 100 (best) | No | 10 | 5 | 0.25–0.90 | 0.0–0.97 | 0.61–2.12 | 5–21 | NR |
| Lysholm ^a | 8 | Likert (3–6) | Patient | 0 (worst) to 100 (best) | Yes | 5 | 3 | 0.65–0.73 | 0.88–0.97 | 0.90–1.10 | 8.9–10.1 | NR |
| OKS ^a | 12 | Likert (5) | Patient | 0 (best) to 48 (worst) | No | 5 | 5 | 0.97–0.93 | 0.91–0.94 | 0.7 | 6.1 | NR |
| CKRS ^b | 13 | Mix | Patient + clinician | 0 (worst) to 100 (best) | Yes | 20 | 10 | NR | 0.75–0.98 | 0.72–2.48 | NR | NR |
| WOMAC ^c | 24 | Likert (5) | Patient | 0 (best) to 20 pain, 8 stiff, 68 funct (worst) | No | 10 | 5 | 0.67–0.98 | 0.65–0.98 | 0.40–2.02 | 10.6–30.6 | 14–33 |
| WOMAC VAS ^a | 24 | VAS/ NRS | Patient | 0 (best) to 500 pain, 200 stiff, 1700 funct (worst) | No | 5 | 3 | NR | NR | NR | NR | NR |
| KSS ^c | 7 | Likert (2–25) | Patient + clinician | 0 (worst) to 100 (best) | Yes | 10 | 5 | 0.74–0.94 | 0.65–0.88 | NR | NR | NR |
| HSS ^d | 13 | Mix | Patient + clinician | 0 (worst) to 100 (best) | Yes | 10 | 5 | 0.70 | 0.98–0.99 | NR | NR | NR |
| AKPS ^e | 13 | Likert (3–5) | Patient | 0 (worst) to 100 (best) | No | 5 | 5 | NR | 0.81 | NR | NR | 8–10 |
| VISA-P ^f | 8 | Mix | Patient | 0 (worst) to 100 (best) | No | 5 | 5 | 0.71–0.73 | 0.74 | NR | NR | 14 |

Note: Subjective IKDC (international knee documentation committee subjective score), Objective IKDC (international knee documentation committee objective score), KOOS (knee injury and osteoarthritis outcome), OKS (Oxford knee score), CKRS(Cincinnati knee rating system), WOMAC (Western Ontario and McMaster Universities index), KSS (knee society score), HSS (hospital for special surgery score), AKPS (Kujala anterior knee pain scale), VISA-P (Victorian Institute of Sport Assessment – Patella)

^aCollins NJ, Misra D, Felson DT, Crossley KM, Roos EM. Measures of knee function. *Arthritis Care Res (Hoboken)*. 2011 Nov;63 Suppl 11:S208–28

^bBarber-Westin SD, Noyes FR, McCloskey JW. Rigorous statistical reliability, validity, and responsiveness testing of the Cincinnati knee rating system in 350 subjects with uninjured, injured, or anterior cruciate ligament-reconstructed knees. *Am J Sports Med*. 1999 Jul-Aug;27(4):402–16

^cHamamoto Y, Ito H, Furu M, Ishikawa M, Azukizawa M, Kuriyama S, Nakamura S, Matsuda S. Cross-cultural adaptation and validation of the Japanese version of the new Knee Society Scoring System for osteoarthritic knee with total knee arthroplasty. *J Orthop Sci*. 2015 Sep;20(5):849–53

^dNarin S, Unver B, Bakurhan S, Bozan O, Karatosun V. Cross-cultural adaptation, reliability and validity of the Turkish version of the Hospital for Special Surgery (HSS) Knee Score. *Acta Orthop Traumatol Turc*. 2014;48(3):241–8

^eCrossley KM, Bennell KL, Cowan SM, Green S. Analysis of outcome measures for persons with patellofemoral pain: which are reliable and valid? *Arch Phys Med Rehabil*. 2004 May;85(5):815–22

^fHernandez-Sanchez S, Hidalgo MD, Gomez A. Responsiveness of the VISA-P scale for patellar tendinopathy in athletes. *Br J Sports Med*. 2014 Mar;48(6):453–7

knee-related quality of life (QoL). It is a complete questionnaire, since it explores all the possible domains of a multitude of possible knee pathologies [99]. However, for this reason, acceptability and reliability could be different based on the patient's age and condition, especially on the sport subscale. It has good correlation with the SF-36 score and the WOMAC. For these reasons and its relative simplicity, the KOOS is used extensively, especially in large-volume registries. Moreover, the individual score for each subscale, rather than an aggregate score, allows for clinical interpretation of different interventions in different dimensions. On the other hand, the KOOS has not been validated for telephone and interview administration, which could limit its applicability due to the need for direct patient involvement [21].

A short version of the KOOS which includes only seven items from the ADL and sport subscales is available as the KOOS-PS (physical function short form), which is shorter, faster, and easier to administer in clinical setting [89].

Conditions: Young and middle-aged patients with post-traumatic OA (undergoing TKA), patient with chondral, ligamentous, or meniscal injuries.

46.6.4 Lysholm Score

The Lysholm score is an eight-item patient-reported scale that evaluates knee symptoms such as limp, locking, swelling, instability, pain, stair climbing, and squatting. It is one of the most commonly used clinical scores for knee evaluation, introduced in 1982 [74]. It is extremely popular and widely used in clinical and research settings. It has a limited floor and ceiling effect, making it useful for tracking improvement of interventions or deterioration over time. Moreover, it has a good correlation with the subjective IKDC, Cincinnati knee ligament score, and the WOMAC. However, its main limitation is that it is clinician derived with no patient input. There are concerns about limited reliability and a lack of definition of MCID. The Lysholm score is

graded as excellent (95–100), good (84–94), fair (65–83), or poor (<65) [21].

Conditions: Ligament injuries and surgery, particularly knee conditions with symptoms of instability, but also meniscal tears, cartilage lesions, patellofemoral pain, and knee osteoarthritis.

46.6.5 Oxford Knee Score (OKS)

The OKS is a 12-item patient-reported score developed for patients undergoing TKR [34]. However, it could be used to evaluate OA and early OA. For these reasons it has a good correlation with WOMAC, KOOS, and SF-36 scores. It is valid, reliable and responsive to change of score, which makes it useful in research settings. However, its development based on knee OA limits its use [21].

Conditions: TKR, OA, rheumatoid arthritis.

46.6.6 Cincinnati Knee Rating System (CKRS)

The CKRS, that has been proposed in 1983 and further modified along the years, is a both patient- and clinician-reported form that is composed of 13 scales assessing symptoms (pain, swelling, and giving way), perception of overall knee condition, daily life function (walking, stairs climbing, and squatting), sport function (running, jumping, and pivoting), sport activity, and occupation. The evaluation is completed with physical examination, functional testing with one-legged hoop exercises, and radiographic measurement of joint narrowing. The overall score, rated from 0 to 100, is obtained combining symptoms (20), functional activities (15), physical examination (25), stability (20), radiographic findings (10), and functional testing (10). Based on the results, it could be graded as excellent (>80), good (55–79), fair (30–54), or poor (<30). It is a comprehensive and rigorous scale, with a good reliability and high responsiveness to detect changes. However, it can be quite time-consuming. Its use

is mostly limited to sports medicine ligamentous and meniscal knee conditions [4].

Conditions: Ligamentous injuries and surgery, meniscal allograft and repair.

46.6.7 Western Ontario and McMaster Universities Index (WOMAC)

The WOMAC is a 24-item patient-reported scale that evaluates three domains, each one with a dedicated subscale: pain, stiffness, and functional activities. It is available both as five-point Likert scale and 100-mm VAS or NRS; therefore, based on the type of scoring, different ratings are obtained for the three subscales. However, the obtained values could be converted to a simple 0–100 scale. The WOMAC is one of the most common scales to evaluate patients with knee OA and is validated in numerous languages. Moreover, it has the advantage of being validated for use in person, over the telephone, or electronically. The individual scores for the three domains, rather than the aggregate value, enhance its interpretation for each domain. However, the presence of items related to uncommon tasks could result in missing data, while the lack of difficult tasks makes it not optimal for more active patients. This scale is optimal for research purpose due to its reliability and ability to detect changes especially after surgical and nonsurgical interventions for knee OA and chondral defects [6, 7, 21].

Conditions: Knee OA, cartilage lesions, and ACL injury.

46.6.8 Knee Society Score (KSS)

The KSS is a seven-item both patient- and clinician-based score that integrates subjective assessment of pain with objective features such as flexion or extension lag, ROM, alignment, and laxity. For this reason, it is limited by low reliability and inter- and intra-observer variations.

It is used mostly for TKA evaluation, and it has a good correlation with SF-36 and the OKS score. Based on the obtained values, it could be interpreted as excellent (80–100), good (70–79), fair (60–69), or poor (<60) [43, 102].

Conditions: Knee OA.

46.6.9 Hospital for Special Surgery Score (HSS)

The HSS is a 13-item scale, both patient- and clinician-reported. It evaluates pain, function, ROM, muscle strength, flexion deformity, instability and alignment. It has similar features to the KSS score and, also, could be graded as excellent (85–100), good (70–84), fair (60–69), or poor (<60). It offers a precise evaluation of knee function but lacks in general quality of life assessment. Therefore, it should be used with other scores capable of depicting patient's general condition [84].

Conditions: Knee OA.

46.6.10 Kujala Anterior Knee Pain Scale (AKPS)

The AKPS, also known as “Kujala score,” is a 13-item patient-reported scale that evaluates the subjective response to six activities, regarded as triggers for anterior knee pain such as walking, running, jumping, climbing stairs, squatting, and sitting [66]. Moreover, the evaluation is integrated with objective basic knee characteristics such as swelling, thigh atrophy, flexion contracture, and patellar abnormal movements. Therefore, it is specifically dedicated to painful conditions of the anterior knee and specifically patellofemoral pathologies. It is a simple and fast questionnaire and has a good correlation with Lysholm, KOOS, and VAS pain; however, it does not distinguish between patients with one episode of patellar dislocation and recurrent instability [27].

Conditions: Anterior knee pain conditions, patellofemoral pathologies, especially instability.

46.6.11 Victorian Institute of Sport Assessment-Patella (VISA-P)

The VISA-P score is an eight-item patient-reported questionnaire composed of a VAS and a Likert portion, which assesses pain during activity or functional tests and sport participation. It has been specifically developed for the measurement of patellar tendon-related conditions. It has a good reliability and repeatability and has a good correlation with the VAS pain. Moreover, since its MCID is available, the VISA-P represents one of the most utilized scores for the assessment of treatments for patellar tendinopathy [48].

Conditions: Patellar tendon disorders.

46.7 Measures of Foot and Ankle Function

An extremely wide range of outcome measures have been developed for the evaluation of the foot and ankle in clinical research; during a 10-year period, 139 different scales have been described, and more recently, between 2012 and 2016, as many as 89 measures have been used in literature for this anatomical region. This incredible variety might be detrimental for an evidence-based decision-making and for comparing clinical results [50, 51].

The basic psychometric characteristics, strength, and weakness of the most common scales for foot and ankle function are described (Table 46.5).

46.7.1 American Orthopaedic Foot and Ankle Society Score (AOFAS Score)

First introduced in 1994, the AOFAS score is the most used outcomes measure tools among clinicians. Four questionnaires are present for different parts of the foot: ankle/hind foot, mid-foot, hallux, and lesser toe; each one is composed of nine items divided into three domains (function, alignment, and pain) and rated on a scale from 0

to 100 [45, 64]. The AOFAS scores are not purely patient-reported since it incorporates both subjective and objective data that requires the clinical assessment. Despite its popularity the AOFAS score has limitations due to lack of validation, high inter-observer variability, and poor correlation with other generic PROMs. For these reasons the AOFAS society itself recommended the usage of more validated and standardized outcome scores [90].

Conditions: These region-specific questionnaires have been used to evaluate patients in a wide variety of foot and ankle pathologies such as arthritis, cartilage defects, soft tissue pathologies, and toe and finger deformities.

46.7.2 American Academy of Orthopaedic Surgeons: Foot and Ankle Model (AAOS-FAM)

The AAOS-FAM was released in 2004, and it is a patient-reported questionnaire composed of 25 items divided into 5 subscales: pain, function, stiffness and swelling, giving way, and shoe comfort. Each answer is measured on a scale of 1–5 or 6 and then calculated; the result is a percentage (0–100) where higher numbers represent better function. This scale is increasing in popularity among surgeons; it has good reliability and repeatability [56, 116].

Conditions: AAOS-FAM can be used to compare clinical outcomes in specific foot and ankle pathologies or surgical methods.

46.7.3 Foot Function Index (FFI)

The FFI was developed in 1991 for senior patients with foot-related pathologies; it was considered specific for foot- and ankle-related conditions secondary to rheumatoid arthritis although there is no specific item for this condition in the questionnaire. It is composed of 23 patient-reported questions that assess foot function in three domains: pain, disability, and activity limitation. It has moderate to high correlation with SF-36; this

Table 46.5 Measures of foot and ankle function

| Questionnaire | Items | Options | Administration | Range | Cutoffs | Time compilation (min) | Time calculation (min) | Internal consistency | Test-retest | SRM | MDC | MCID |
|-----------------------|-------|--------------|---------------------|---------------------------|---------|------------------------|------------------------|-------------------------------------|-------------|------|-----|---------------|
| AOFAS ^a | 9 | Likert (3–4) | Patient + clinician | 0 (worst) to 100 (best) | No | 10 | 3 | 0.585 (0.863 for function subscale) | 0.89–0.97 | 0.79 | 1.7 | NR |
| AAOS-FAM ^b | 25 | Likert (5–6) | Patient | 0% (worst) to 100% (best) | No | 15 | 10 | 0.83–0.93 | 0.79–0.99 | NR | NR | NR |
| FFI ^c | 23 | VAS | Patient | 0 (best) to 230 (worst) | No | 5 | 7 | 0.73–0.96 | 0.70–0.87 | NR | 7 | NR |
| FAOS ^b | 42 | Likert (5) | Patient | 0 (worst) to 100 (best) | No | 10 | 10 | 0.88–0.94 | 0.70–0.87 | NR | NR | NR |
| FAAM ^b | 29 | Mix | Patient | 0% (worst) to 100% (best) | No | 7 | 10 | 0.96–0.98 | 0.87–0.89 | NR | NR | 8 ADL/9 sport |
| FADI ^b | 34 | Likert (5) | Patient | 0% (worst) to 100% (best) | No | 7 | 10 | 0.84–0.89 | NR | NR | NR | NR |
| ACFAS ^c | 16 | Likert (5) | Patient + clinician | 0 (worst) to 100 (best) | No | 15 | 10 | NR | NR | NR | NR | NR |
| FHSQ ^b | 13 | Likert (3–5) | Patient | 0 (worst) to 100 (best) | No | 6 | 10 | 0.89–0.95 | 0.74–0.92 | NR | NR | 7–14 |
| ROFPAQ ^b | 39 | Likert (5) | Patient | 1 (best) to 5 (worst) | No | 10 | 6 | 0.81–0.89 | 0.82–0.93 | NR | NR | NR |
| QOL ^b | 5 | Likert (5) | Patient | 0 (worst) to 20 (best) | No | 3 | 1 | 0.85–0.91 | NR | NR | NR | NR |
| OMAS ^d | 9 | Likert (3–5) | Patient | 0 (worst) to 100 (best) | No | 3 | 3 | 0.76 | 0.95 | NR | NR | NR |
| VISA-A ^e | 8 | Mix | Patient | 0 (worst) to 100 (best) | No | 5 | 7 | 0.73 | 0.79 | NR | NR | NR |

Note: AOFAS (American orthopaedic foot and ankle society score), AAOS-FAM (American Academy of Orthopaedic: foot and ankle), FFI (foot function index), FAOS (foot and ankle outcomes score), FAAM (foot and ankle ability measure), FADI (foot and ankle disability index), ACFAS (American College of Foot and Ankle Surgeons), FHSQ (foot health status questionnaire), ROFPAQ (Rowan foot pain assessment), QOL (sport ankle QOL), OMAS (Olerud-Molander ankle score), VISA-A (Victorian Institute of sports assessment-Achilles)

^aDe Boer AS, Meuffels DE, et al. Validation of the American Orthopaedic Foot and Ankle Society Ankle-Hindfoot Scale Dutch language version in patients with hindfoot fractures. *BMJ Open*. 2017;7(11):e018314

^bMartin RL, Irrgang JJ. A survey of self-reported outcome instruments for the foot and ankle. *J Orthop Sports Phys Ther*. 2007;37(2):72–84

^cCook JJ, Cook EA, et al. Validation of the American College of Foot and Ankle Surgeons Scoring Scales. *J Foot Ankle Surg*. 2011;50(4):420–9

^dNilsson et al. The Swedish version of OMAS is a reliable and valid outcome measure for patients with ankle fractures. *BMC Musculoskeletal Disorders*. 2013;14:109

^eIversen, J. V., Bartels, et al. Danish VISA-A questionnaire. *Scand J Med Sci Sports*. 2016; 26: 1423–1427

suggests that FFI may be a good measure of both health status and patients' outcomes [13, 14, 104].

Conditions: Generally used in older patients, rheumatoid patients, orthotics outcomes, poor reliability in professional athlete due to the reported ceiling effect.

46.7.4 Foot and Ankle Outcomes Score (FAOS)

The FAOS, released in 2001, is a 42-item patient-reported outcomes measure that consists in five subscales (pain, symptoms, ADL, sport, and ankle-related quality of life). Each subscale is graded separately and scored in a 0–100 value. The FAOS demonstrated good reliability and validity, but the length of this survey can create significant burden for the patient [41, 98].

Conditions: It has been validated for a variety of foot and ankle pathologies such as adult flat-foot deformity, hallux valgus, hallux rigidus.

46.7.5 Foot and Ankle Ability Measure (FAAM)

The FAAM was developed in 2005; it is region-specific and composed of 29 patient-reported items divided in activity of daily living (ADL) and sport subscales. A recent study demonstrated that the FFI and FAAM are highly correlated for foot and ankle trauma patients [40, 77].

Conditions: It is valid for a range of foot and ankle conditions as well as for chronic ankle instability and diabetes mellitus-related conditions.

46.7.6 Foot and Ankle Disability Index (FADI)

The FADI was first released in 1999; it is the former version of FAAM and includes four more items for pain assessment and one item for the ability to sleep (34 total items). FADI and FAAM are appropriate to evaluate functional disabilities in athletes with chronic ankle instability [76].

Conditions: Sport-related foot and ankle pathologies and trauma evaluation.

46.7.7 American College of Foot and Ankle Surgeons (ACFAS) Universal Evaluation Scoring Scales

The American College of Foot and Ankle Surgeons developed these anatomically based scoring scales in 2005 as clinical instruments to evaluate objective and subjective parameters before and after surgery [106]. Four modules exist for the first metatarsal-phalangeal joint and first ray, the forefoot (excluded first ray), the rear foot, and the ankle; each questionnaire is completed by both patient and clinician and includes subjective (pain, appearance, and functional capacity) and objective (radiographic and functional) parameters, for a total of 100 points. This instrument has been validated and presents good reliability and sensitivity to change [24].

Conditions: Foot and ankle musculoskeletal-related pathologies requiring surgical intervention.

46.7.8 Foot Health Status Questionnaire (FHSQ)

The FHSQ was developed for individuals undergoing surgical treatment for common foot conditions. It consists of four subscales with a total of 13 items representing the following four domains: pain (four items), function (four items), footwear (three items), and general foot health (two items). Scores from each subscale range from 0 to 100, with a higher score representing better outcomes [8].

Conditions: Foot- and ankle-related disorders including those affecting skin and nail.

46.7.9 Rowan Foot Pain Assessment (ROFPAQ)

The ROFPAQ was developed as a disease-specific instrument for chronic foot pain. It contains

39 items in the following four subscales for pain assessment: sensory (16 items), affective (ten items), cognitive (ten items), and comprehension (three items). Each subscale is scored independently from 1 through 5, and the item responses are merged together to produce a subscale score ranging from 1 to 5, with a higher score representing more pain [100].

Conditions: Chronic foot and ankle pain.

46.7.10 Sport Ankle QOL (Quality of Life)

The sport ankle rating system quality of life measure was developed as a region-specific measure including self-reported and clinician-completed outcome measures. The QOL measure, the clinical rating score, and a single numeric evaluation are the three outcome measures that could be used together or independently. The QOL is a self-reported questionnaire designed to assess an athlete's quality of life after an ankle injury; it contains five items that evaluate symptoms, work and school activities, recreation and sports activities, activities of daily living, and lifestyle.

The clinical rating score is composed of 11 items both patient and clinician based; finally, with the numeric VAS evaluation, the patient is asked to score his ankle function from 0 to 100 [113].

Conditions: Ankle injuries and specifically ankle sprains.

46.7.11 Olerud-Molander Ankle Score (OMAS)

The OMAS is a disease-specific outcomes measure developed for patients with ankle fractures and has been frequently used to evaluate this group of subjects; furthermore, it has been reported to be a valid item for recording short-term changes after an acute ankle ligament injury. OMAS is a self-administered patient questionnaire; the scale ranges from 0 points (totally impaired function) to 100 points (completely unimpaired function) and is based on nine differ-

ent domains: pain, stiffness, swelling, stair climbing, running, jumping, squatting, supports, and work/activity level [87, 88].

Conditions: Ankle fracture, ligament ankle injury.

46.7.12 Victorian Institute of Sports Assessment-Achilles (VISA-A)

The VISA-A is a disease-specific instrument designed to evaluate the clinical severity for patients with chronic Achilles tendinopathy. It is an easily self-administered questionnaire that evaluates symptoms and their effect on physical activity. The questionnaire contains eight questions, covering three necessary domains: pain, functional status, and activity. The first six questions use a visual analog scale so that the patient may report the magnitude of a continuum of subjective symptoms; the final two questions used a categorical rating scale. The final results range from 0 to 100, with asymptomatic persons expected to score 100 points [95].

Conditions: Chronic Achilles tendinopathy.

46.8 Measures of Activity Level

Making patients capable of an unlimited physical activity is the main focus of clinicians; for this reason several scores have been created to assess outcomes in terms of return to sport/activity (RTS). While considering these instruments, a factor to be outlined is that athletes are different from the general population since they have higher level of physical function and perceived health, often they do not perceive symptoms during the daily activities, and common outcome measures may not detect problems that only result from high-intensity training and competition.

The basic psychometric characteristics, strength, and weakness of the most common scales for activity level are described below (Table 46.6).

Table 46.6 Measures of activity level

| Questionnaire | Items | Options | Administration | Range | Cutoffs | Time compilation (min) | Time calculation (min) | Internal consistency | Test-retest | SRM | MDC | MCID |
|----------------------|----------------|--------------|----------------|--|---------|------------------------|------------------------|----------------------|-----------------------|-----|---------------------|------|
| TEGNER ^a | 1 | Likert (11) | Patient | 0 (worst) to 10 (best) | No | 1 | 1 | NR | 0.8 | 1.0 | 1 | 1 |
| UCLA ^b | 1 | Likert (10) | Patient | 1 (worst) to 10 (best) | Yes | 2 | 2 | NR | NR | NR | 1 | 1 |
| ARS ^c | 4 | Likert (5) | Patient | 0 (worst) to 16 (best) | No | 1 | 1 | 8.87 | 0.81 | NR | NR | NR |
| AAS ^d | 1 | Likert (11) | Patient | 0 (worst) to 10 (best) | No | 1 | 1 | 1.0 | NR | NR | 1 | 1 |
| LEFS ^e | 20 | Likert (5) | Patient | 0 (worst) to 80 (best) | No | 5 | 3 | 0.96 | 0.86 | NR | 9 | 9 |
| SAS ^f | 7 | Likert (4–5) | Patient | 0 (worst) to 20(best)/A–D | No | 1 | 1 | NR | 0.92 | NR | NR | NR |
| HAS ^g | 44 | Likert (6) | Patient | 0 (worst) to 220 (best) | No | 120 | 30 | NR | 0.75 | NR | NR | NR |
| OSTRC ^h | 4 (each joint) | Likert (4–5) | Patient | 0 (best) to 100 (worst) | No | >5 (variable) | 3–7 | 0.86–0.91 | NR | NR | NR | NR |
| SQUASH ⁱ | 11 | Open | Patient | 1 min. to 9 max. (per item) no upper limit | No | 3–5 | 10 | NR | NR | NR | NR | NR |
| IPAQ-SF ^j | 7 | Open | Patient | 0 to No upper limit | No | 3–5 | 10 | <0.8 | NR | NR | NR | NR |
| HAP ^k | 94 | Likert (3) | Patient | 1 (worst) to 94 (best) | No | 10 | 5 | NR | 0.84 MAS; 0.79 AAS | NR | 6.5 MAS; 8.4 AAS | NR |

Note: ARS (activity rating scale), AAS (ankle activity score), LEFS (lower extremity functional scale), SAS (shoulder activity scale), HAS (Heidelberg sport activity score), OSTRC (Oslo Sports Trauma Research Centre overuse injury questionnaire), SQUASH (short questionnaire to assess health-enhancing physical activity), IPAQ-SH (international physical activity questionnaire – short form), HAP (human activity profile)

^aBriggs KK, Lysholm J, et al. The reliability, validity, and responsiveness of the Lysholm score and Tegner activity scale for anterior cruciate ligament injuries of the knee: 25 years later. *Am J Sports Med.* 2009;37(5):890–7

^bTerwee CBI, Bouwmeester W, et al. Instruments to assess physical activity in patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. *Osteoarthritis Cartilage.* 2011;19(6):620–33

^cHosseini Negahban, Neda Mostafaei, Soheil Mansour Sohani, Masood Mazaheri, Shahin Goharpey, Mahyar Salavati, Shabla Zahednejad, Zohreh Meshkati & Ali Montazeri (2011) Reliability and validity of the Tegner and Marx activity rating scales in Iranian patients with anterior cruciate ligament injury, *Disability and Rehabilitation*, 33:22–23, 2305–2310

^dTamás Halasi, MD, Ákos Kynsburg, MD, et al. Development of a New Activity Score for the Evaluation of Ankle Instability. *The American Journal of Sports Medicine.* 2004;32:899–908

^eAlcock GK, Stratford PW. Validation of the Lower Extremity Function Scale on athletic subjects with ankle sprains. *Physiother Can.* 2002;54:233–240

^fBrophy RH, Beauvais RL, et al. Measurement of shoulder activity level. *Clin Orthop Relat Res.* 2005;439:101–8

^gSeeger J, Weinmann S, et al. The Heidelberg Sports Activity Score - A New Instrument to Evaluate Sports Activity. *The Open Orthopaedics Journal.* 2013;7:25–32

^hClarsen B, Myklebust G, et al. Development and validation of a new method for the registration of overuse injuries in sports injury epidemiology: the Oslo Sports Trauma Research Centre (OSTRC) Overuse Injury Questionnaire. *Br J Sports Med.* 2013;47:495–502

ⁱTerwee CB, Bouwmeester W, et al. Instruments to assess physical activity in patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. *Osteoarthritis Cartilage.* 2011;19(6):620–33

46.8.1 Tegner Activity Score

First described in 1985 [105] for the prospective evaluation of the knee ligaments injuries, the Tegner activity scale provides an arbitrary ranking based on the level of sport and leisure time activities and competition at which the patient is currently participating. It is a simple scale in which the subject indicates his/her current activity ranging from 0 (no physical activity/disabled) to 10 (participation in competitive soccer or pivoting sports). It was created as a complement to the Lysholm score; but its use has also extended into other joints, including the hip and ankle [42, 78].

Conditions: Tegner score was developed and is mostly used for knee ligamentous injuries and reconstructions.

46.8.2 University of California at Los Angeles (UCLA) Activity Rating Scale

The UCLA activity rating scale is a simple scale ranging from 1 (no activity) to 10 (participation in impact sports); it was developed in 1998 to assess physical activity after joint replacements. Like the Tegner score, the patient is asked to rate his/her own most appropriate activity level. Four activity subgroups were defined: scores between 0 and 4 (low activity), 4.1 and 6 (moderately low activity), 6.1 and 8 (moderately high activity), and 8.1 and 10 (high activity) [114].

Conditions: The UCLA is mostly used and validated for hip and knee osteoarthritis and evaluation of joint replacement.

46.8.3 Activity Rating Scale (ARS) or Marx Scale

The ARS/Marx questionnaire quantifies the frequency of activities that challenge the dynamic stability of the knee; it consists in four questions about how frequently the patients perform activity such as running, cutting, decelerating, and pivoting. Each question is scored from 0 (<1

time/month) to 4 (>4 time/month), and the total score range is 0–16. ARS is based on the idea of measuring specific components of function/movement (that apply universally to the lower limb) to allow more accurate comparison among patients. This scale can be completed in a very short time frame [78].

Conditions: Sport activity involving complex articular motions of the knee and lower limb.

46.8.4 Ankle Activity Score (AAS)

The AAS is a joint-specific score that was published in 2004; it was based on the Tegner score. It contains 53 sports, three working activities, and four general activities; the patient is asked to select his/her most appropriate sport/activity and to indicate a level of participation (top level, lower competitive level, recreational level). The result, as with the Tegner score, is represented by a single number from 0 to 10 [42].

Conditions: Ankle injuries.

46.8.5 Lower Extremity Functional Scale (LEFS)

The LEFS was developed to be a broad region-specific measure for individuals with musculoskeletal disorders of the hip, knee, ankle, or foot. It consists of 20 items that specifically cover the domains of activity and participation. The scale uses a Likert response format, with a higher score representing a higher level of ability [10].

Conditions: LEFS have been validated for several pathologies of the lower limb; moreover, it has been translated in different languages.

46.8.6 Shoulder Activity Scale (SAS)

The Brophy-Marx SAS was developed in 2005 as an easy instrument to evaluate the patient's overall shoulder activity level that could be generalized across different sports and completed in less than 1 min. It is composed of two parts: the first five items describe five common activities of the

shoulder and the relative frequency, during the patient's previous year, for each item is scored from 0 to 4 (never or less than once a month, once a month, once a week, more than once a week, or daily). The total numerical activity score ranges from a minimum of 0 points to a maximum of 20 points. In the second part of the score, the patients are asked if they participate in contact sports and sports that involve repetitive overhead throwing. The answers of these two questions range from A (no) to D (yes, at professional level). SAS has shown good reliability and validity [11, 12].

Conditions: This score has been developed on patients with rotator cuff tears.

46.8.7 Heidelberg Sport Activity Score (HSAS)

The HAS was published in 2013; this validated instrument divides sport activities into 11 categories: walking, swimming, cycling, running, cross-country skiing, alpine skiing, golfing, dancing, racket sports, ball sports, and miscellaneous. For each of these activities, the patient is asked to grade between 0 and 5 about frequency, duration, level of importance, and impairment from the affected joint. For each activity, a core from 0 to 20 is calculated with a formula: $(\text{frequency} + \text{duration}) \times (1 + \text{impairment}/10 + \text{importance}/10)$. The scores are then added to obtain a final score between 0 and 220. HAS has proven high validity, sensitivity, reliability, and sensitivity. It can be used for elite-level athletes and athletes who perform different sports and is valid for different joints; nevertheless, its disadvantage is an extremely long time for compilation (120 min) [103].

Conditions: Evaluation of activity after trauma or surgery; it can be used in elite-level athletes.

46.8.8 Oslo Sports Trauma Research Center Overuse Injury Questionnaire (OSTRC)

The intention of the OSTRC was to create a questionnaire that could be applied to overuse

injury problems in any area of the body. The instrument is designed in four items for each affected joint; the final "severity score" ranges between 0 and 100 (25 point for item) for each overuse problem. In studies with multiple anatomical areas of interest, the four questions are repeated for each area. This questionnaire uses the term "problem" rather than "injury" since there is greater variation in interpretation of the term "injury" [20].

Conditions: OSTRC is mainly used for the evaluation of overuse problems in sports injury epidemiology.

46.8.9 Short Questionnaire to Assess Health-Enhancing Physical Activity (SQUASH)

The SQUASH was not designed to measure energy expenditure but to give an indication of the habitual activity level. It consists of 11 questions on commuting activities, leisure time and sports activities, household activities, and activities at work and school. The total activity score is calculated by taking the sum of the activity scores for separate questions [112].

Conditions: SQUASH is a short physical activity questionnaire with the general purpose to assess habitual physical activity.

46.8.10 International Physical Activity Questionnaire-Short Form (IPAQ-SF)

The IPAQ-SF consists in seven questions about the frequency and durations of participation in strenuous, moderate, and walking activities in addition to the time spent sitting during the past week. The final score is expressed in metabolic equivalents (METs) which represent the oxygen consumption of an individual sitting for 1 min (3.5 mL/kg/min) [26, 29].

Conditions: IPAQ has been validated, and it presents reasonable measurement properties for monitoring population levels of physical activity in diverse settings.

46.8.11 Human Activity Profile (HAP)

The HAP is a 94-item self-report measure of energy expenditure or physical fitness; it was developed as an outcome measure for medical rehabilitation for people with a wide spectrum of physical disorders. It consists of a list of activities for which patients should indicate if they are currently able to perform the activity, have stopped performing the activity, or have never performed the activity. Each of the selected activities has an estimate energy requirement between approximately 1 and 10 METs. Two scores are calculated: the maximum activity score (MAS) and the adjusted activity score (AAS) [38].

Conditions: Epidemiologic and population studies as well as rehabilitation medicine.

Fact Box 46.2

Athletes population is in many ways different from general one; they have much higher level of physical functioning and health status. For this reason, they do not perceive symptoms during the daily activities and choosing an adequate outcome measure is compulsory.

46.9 Measures of Global and Mental Health

Generic measures of health-related quality of life are frequently used to evaluate the impact of treatments and clinical results and to monitor population health. Often these scales are composed of various independent domains/dimensions that together represent the notion of health-related quality of life. The items are weighted to indicate the relative importance attributed to them by the respondents and then aggregated into a single number reflecting the quality or value of a health state. To obtain such values, several instruments have been developed.

The basic psychometric characteristics, strength, and weakness of the most common scales for global and mental health are described (Table 46.7).

46.9.1 36-Item Short-Form Health Survey (SF-36) and Short-Form 12 (SF-12)

The SF-36 is a general health measure, introduced in 1992 that includes 36 items addressing eight domains of overall health status: physical functioning (PF), bodily pain (BP), role limitations due to physical health problems (RP), role limitations due to personal or emotional problems (RE), general mental health (MH), social functioning (SF), energy/fatigue or vitality (VIT), and general health perceptions (GH). Although this scale has been validated for orthopedic use, experts recommend pairing the SF-36 with an orthopedic-specific measure since it is a general health scale, and it could be difficult to isolate orthopedic outcomes from other unrelated health conditions.

The SF-12 is a shortened version of the SF-36, developed in 1996, with the aim of reducing redundancies and time burden on the patient. It shortens the survey to 12 items and reports 2 scores in physical and mental domains. It has been validated for orthopedic patients. The SF-12 was included as a recommended PRO measure for “general quality of life” by the AAOS.

Both SF-36 and SF-12 are great questionnaires for outcomes assessment in research; both need to be administered in conjunction with orthopedic-specific measures [15, 52, 110, 111, 115].

46.9.2 EuroQol-5 Domains-3 Likert (EQ-5D-3L)

The EQ-5D health status and quality-of-life measure is composed of five items (mobility, self-care, usual activity, pain/discomfort, and anxiety/depression), with three possible response levels (no problems, some/moderate problems, extreme problems). The EQ-5D index is

Table 46.7 General and mental health measures

| Questionnaire | Items | Options | Administration | Range | Cutoffs | Time compilation (min) | Time calculation (min) | Internal consistency | Test-retest | SRM | MDC | MCID |
|------------------------|-------|--------------|----------------|----------------------------------|---------|------------------------|------------------------|------------------------------------|--------------------|----------|-----|------|
| SF-36 ^a | 36 | Likert (2–6) | Patient | 0 (worst) to 100 (best) | No | 7 | 13 | ≥0.70 (0.90 for physical function) | <0.70 | 1.04 | 5 | NR |
| SF-12 ^a | 12 | Likert (2–6) | Patient | 0 (worst) to 100 (best) | No | 3 | 7 | ≥0.82 PCS; ≥0.75 MCS | 0.89 PCS; 0.79 MCS | <0.2 | NR | NR |
| EQ-5D-3L ^b | 6 | Mix | Patient | −0.594 (worst) to 1 (best)/0–100 | No | 5 | 10 | NR | NR | 0.7/0.04 | NR | NR |
| AQoL ^c | 12 | Likert (4) | Patient | −0.04 (worst) to 1.0 (best) | No | 5 | 7 | NR | NR | NR | NR | NR |
| NPH ^a | 45 | Dichotomous | Patient | 0 (best) to 100 (worst) | No | 5–10 | 10–15 | 0.83 (0.33 for mobility) | 0.77–0.85 | NR | NR | NR |
| PROMIS 10 ^b | 10 | Mix | Patient | t-score distribution | No | 5 | 2 | NR | NR | 0.72 | NR | NR |

Note: SF-36 (36-Item Short Form Health Survey), SF-12 (12-Item Short Form Health Survey), EQ-5D-3L (EuroQoL-5 Domains-3 Likert), AQoL (Assessment of Quality of Life), NPH (Nottingham Health Profile), PROMIS (Patient-Reported Outcomes Measurement Information System 10 Global health)

^aBusija, L., Pausenberger, E., et al. (2011), Adult measures of general health and health-related quality of life: Medical Outcomes Study Short Form 36-Item (SF-36) and Short Form 12-Item (SF-12) Health Surveys, Nottingham Health Profile (NHP), Sickness Impact Profile (SIP), Medical Outcomes Study Short Form 6D (SF-6D), Health Utilities Index Mark 3 (HUI3), Quality of Well-Being Scale (QWB), and Assessment of Quality of Life (AQOL). *Arthritis Care Res*, 63: S383-S412

^bOak SR, Strnad GJ, et al. Responsiveness Comparison of the EQ-5D, PROMIS Global Health, and VR-12 Questionnaires in Knee Arthroscopy. *Orthop J Sports Med*. 2016 Dec 17;4(12):2325967116674714

^cKathryn Whitfield, Rachel Buchbinder, et al. Parsimonious and efficient assessment of health-related quality of life in osteoarthritis research: validation of the Assessment of Quality of Life (AQoL) instrument. *Health and Quality of Life Outcomes* 2006, 4:19

calculated from the five dimensions, ranging from -0.594 (worst) to 1.0 . Moreover, to the EQ-5D index, the EQ-5D includes a VAS for rating of overall health status from 0 (worst imaginable health) to 100 (best imaginable health). A common criticism of this measure is the lack of sensitivity to change since only three levels of responses are available within each construct. With the aim of addressing this issue, a version of the measure with five responses has been developed, called the EQ-5D-5L [47].

46.9.3 Assessment of Quality of Life (AQoL)

The AQoL is a 12-item instrument which loads onto four dimensions: independent living, social relationships, physical senses, and psychological well-being. These subscales are weighted between 0.0 (death) and 1.0 (full health). With its emphasis upon psychosocial dimensions of health, it offers significant advantages for evaluation studies where these dimensions are important [93].

46.9.4 Nottingham Health Profile (NPH)

The NHP questionnaire is a self-administered questionnaire. It was developed in English and consists of two parts: Part 1 contains 38 “yes/no” questions covering six dimensions: pain, physical mobility, emotional reactions, energy, social isolation, and sleep. Part 2 has seven “yes/no” questions concerning problems of daily activities. It has been shown to be internally consistent, valid, reproducible, and sensitive [55].

46.9.5 Patient-Reported Outcomes Measurement Information System 10 Global Health (PROMIS-10 Global Health)

The PROMIS was established in 2004 with funding from the National Institutes of Health; this

initiative develops and evaluates standard measures for key patient-reported health indicators and symptoms. PROMIS measures are standardized, allowing for assessment of many patient-reported outcome domains such as pain, fatigue, emotional distress, physical functioning, and social role participation.

Computerized adaptive testing (CAT) software has been implemented; this allows tailoring the PRO assessment to the individual patient by selecting the most informative set of questions based on responses to previous questions [17].

46.10 Measures of Pain

Pain is a complex and subjective experience and that implies several measurement challenges. It is important for the clinicians to utilize sensitive and accurate pain outcome measures although currently we rely mainly on self-report measures. The cutoff value for clinical significance of pain reduction must be determined on the minimal amount of change being important to patients. A reduction of 10–20% of pain can be considered clinically significant [67].

The basic psychometric characteristics, strength, and weakness of the most common scales for pain assessment are described (Table 46.8).

46.10.1 Visual Analog Scale for Pain (VAS for Pain)

The VAS for pain was introduced in 1976; it is a widely recognized and simple instrument that allows the patient to score his own pain level on a straight 100-mm line with zero indicating “no pain” and 100 “worst imaginable pain.” Its usefulness in orthopedic surgery has been recognized, and VAS has proven high validity and responsiveness; on the other hand, its low specificity has been shown with a 1.1 cm decrease corresponding to the minimal clinical important difference for pain. The patient acceptable symptomatic state is considered with a value less than 3 cm [37, 46, 59, 115].

Table 46.8 Pain measures

| Questionnaire | Items | Options | Administration | Range | Cutoffs | Time compilation (min) | Time calculation (min) | Internal consistency | Test-retest | SRM | MDC | MCID |
|---------------------|-------|-------------|----------------|--------------------------------|---------|------------------------|------------------------|----------------------|-------------|-------|----------------------------|--------|
| VAS ^a | 1 | VAS | Patient | 0 mm (best) to 100 mm (worst) | Yes | <1 | <1 | NR | 0.71–0.94 | NR | 1 mm | 1.1 cm |
| NRS ^a | 1 | Likert (11) | Patient | 0 (best) to 10 (worst) | No | <1 | <1 | NR | 0.96 | NR | 1 | 2 |
| VRS ^a | 1 | Likert (5) | Patient | 1 (best) to 5 (worst) | No | <1 | <1 | NR | NR | NR | 1 | 1 |
| FPS-R ^b | 1 | Likert (6) | Patient | 0 (best) to 10 (worst) | No | <1 | <1 | NR | NR | NR | 2 | 2 |
| SF-MPQ ^a | 17 | Mix | Patient | 0–45/0–5/0–10 (best to (worst) | No | 2–5 | 1–2 | 0.73–0.89 | 0.79–0.93 | >0.80 | 5.2, 4.5, 2.8, 1.4, 1.4 cm | >5 |

Note: VAS (visual analog scale for pain), NRS (numerical rating scale for pain), VRS (verbal rating scale for pain), FPS-R (faces pain scale revised), SF-MPQ (short-form McGill pain questionnaire)

^aHawker, G. A., Mian, S., et al. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care Res*, 2011;63: S240-S252

^bFerreira-Valente MA, Pais-Ribeiro JL, et al. Validity of four pain intensity rating scales. *Pain*. 2011;152(10):2399–404

46.10.2 Numerical Rating Scale for Pain (NRS for Pain)

The NRS is an 11-point scale consisting of integers from 0 through 10: 0 representing “no pain” and 10 representing “worst imaginable pain.” Respondents select the single number that best represents their pain intensity. It is considered to be more comprehensive compared to the VAS for; however, it may capture the complex nature of the pain experience [46].

46.10.3 Verbal Rating Scale for Pain (VRS for Pain)

The VRS is a single domain five-point scale consisting of a list of sentences (no pain, mild pain, moderate pain, intense pain, maximum pain) describing increasing levels of pain severity. Respondents select the single phrase that best characterizes their pain intensity [46].

46.10.4 Faces Pain Scale-Revised (FPS-R)

The FPS-R is a six-point scale represented by six different faces showing increasing severity of pain. Patients are asked to select the facial expression that best resembles his or her pain intensity, from the left-most face (“no pain”) to the right-most face (“very much pain”). The FPS-R was originally developed for pediatric patients, but its simplicity makes it a reliable instrument for individuals with cognitive and communication impairments as well [9, 49, 101].

46.10.5 Short-Form McGill Pain Questionnaire (SF-MPQ)

The SF-MPQ is a multidimensional measure, with extensive clinical research use. Patients rate their pain in sensory terms (e.g., sharp or stabbing) and affective terms (e.g., sickening or fearful), with 15 total descriptors. Each item is rated on a four-point scale that ranges from none to

severe. The SF-MPQ also has a single VAS item for pain intensity and a VRS for rating the overall pain experience. It is used particularly to measure the sensory and affective aspects of pain and pain intensity in adults with chronic pathologies [46, 79].

Fact Box 46.3

Since “pain” perception is maybe the most relevant outcome in clinical practice, research protocols should include sensitive and accurate pain measures. Currently we rely mainly on self-report measures where a reduction of 10–20% of pain can be considered as the minimal amount of change being important to patients.

46.11 Measures of Sport-Related Psychological Aspects

It has been demonstrated that while most athletes reach a normal physical function, less than half of them return to the same level of sport activity. Possibly, psychological factors are involved in the rehabilitation processes and in the athlete’s self-perception of recovery. The following section describes some of the common scales for sport activity assessment, energy expenditure, and psychological factors after injuries.

The basic psychometric characteristics, strength, and weakness of the most common scales for sport-related psychological aspects are described (Table 46.9).

46.11.1 Injury-Psychological Readiness to Return to Sport Scale (I-PRRS)

The I-PRRS is an easy to use tool developed to measure the athlete’s psychological readiness to return to sport after injury. It is a six-item scale, each item is scored from 0 (no confidence) to 100 (maximum confidence) with intervals of 10. The scores from the six items are summed and divided

Table 46.9 Psychological scores and measures

| Questionnaire | Items | Options | Administration | Range | Cutoffs | Time compilation (min) | Time calculation (min) | Internal consistency | Test-retest | SRM | MDC | MCID |
|--------------------|-------|-------------|----------------|-------------------------|---------|------------------------|------------------------|----------------------|-------------|-----|-----|------|
| IPRRS ^a | 6 | Likert (11) | Patient | 0 (worst) to 60 (best) | Yes | 8–10 | 5–8 | 0.63 | 0.97 | NR | NR | NR |
| RIAI ^b | 28 | Likert (4) | Patient | 0 (best) to 45 (worst) | No | 8–15 | 8 | 0.96–0.98 | NR | NR | NR | NR |
| TSK ^c | 11 | Likert (4) | Patient | 11 (best) to 44 (worst) | No | 4–8 | 5 | 0.7–0.9 | >0.7 | NR | NR | NR |

Note: IPRRS (injury-psychological readiness to return to sport scale), RIAI (re-injury anxiety inventor), TSK (Tampa scale for kinesiophobia)

^aNaghdi S, Nakhostin Ansari N, et al. Cross-cultural adaptation and validation of the Injury-Psychological Readiness to Return to Sport scale to Persian language. *Physiother Theory Pract.* 2016;32(7):528–35

^bWalker, Natalie et al. A preliminary development of the Re-Injury Anxiety Inventory (RIAI). *Physical Therapy in Sport*, Volume 11, Issue 1, 23–29

^cRoelofs J, van Breukelen G, et al. Norming of the Tampa Scale for Kinesiophobia across pain diagnoses and various countries. *Pain.* 2011;152:1090–1095

by 10 to calculate the total score. The range of scores is between 0 and a maximum score of 60. A score of 60 indicates high confidence to return to sport; 40, moderate confidence; and 20, low confidence [39].

Conditions: Evaluation of psychological readiness to return to sport among athletes.

46.11.2 Re-injury Anxiety Inventory (RIAI)

The RIAI is a 28-item score designed to assess the athlete's fear of experiencing a re-injury. It is composed of the rehabilitation anxiety (RIA-R, 13 items) and the reentry to competition anxiety (RIA-RE, 15 items). The instrument is based on a four-point (0–3) Likert-response type; the final score ranges from 0 (complete absence of anxiety) to 45 (extreme anxiety) [109].

Conditions: RIAI can be used in studies aiming to evaluate athlete's psychological readiness for RTS.

46.11.3 Tampa Scale for Kinesiophobia (TSK)

The TSK is a self-reported measure developed to assess "fear of movement-related pain" in patients with musculoskeletal disorders. The original test, developed in English 1, has been translated into ten languages. The TSK-11 is the most widely used; it contains 11 items from the original 17-item questionnaire. Each item is scored on a four-point Likert scale, ranging from 1 "strongly disagree" to 4 "strongly agree"; total scores vary between 11 and 44, with higher scores indicating higher levels of fear of movement-related pain [96].

Conditions: The TSK-11 is a reliable and valid measurement tool that provides therapists valuable information on activity avoidance and pathological somatic focus in patients with musculoskeletal pain.

Clinical Vignette

An innovative and minimally invasive surgical technique for complex osteochondral knee lesions was developed in your institution. From the first outpatient follow-ups, you realize that the subjects treated with this new technique seem to be very happy about their health status and knee function. Finally, you are charged to design a study protocol to compare the result of the new technique with the classic one. Beside an accurate imaging of the bone and cartilage and maybe a biochemical characterization of the fluids, which clinical outcomes measures can be included in our protocol? Which ones are more indicated to detect an effective improvement in patient conditions?

First one or more knee specific measures should be chosen. In this pathology, the KOOS score has demonstrated good psychometric properties, and the WOMAC score has an excellent reliability and ability to detect changes.

Secondly, we want to assess the patient's perceived pain level and health status. For this purpose, the SF-MPQ score for pain is accurate and easy to complete; moreover it includes a VAS for general pain assessment, while the SF-12 score has a short-time compilation and will give a precise overview of patient's general health.

Finally, since most of our patients used to be quite active before injury, our protocol should include at least one activity level measure. We choose the Tegner score, which was specifically designed for knee injuries and is extremely intuitive in its compilation, and the LEFS score which can be used for several lower limb pathologies and shows good reliability.

Take-Home Message

- The development, testing, and implementation of tools to aid in the measurement of phenomena in medicine are central to clinical practice and clinical research; therefore PROMs are a key component to orthopedics research and may also be so for clinical practice in orthopedic surgery.
- Health status measurement instruments must possess adequate measurement properties, and it is fundamental to remember that a complete picture of a condition or a treatment effect on a patient could be provided only with the combination of a “disease-specific measure” and a “generic measure.”

References

1. Aichroth PM, Cannon WD Jr. International Knee Documentation Committee: knee ligament injury and reconstruction 209 evaluation. Knee surgery current practice. New York: Raven Press; 1992. p. 759–60.
2. Amstutz HC, Sew Hoy AL, Clarke IC. UCLA anatomic total shoulder arthroplasty. *Clin Orthop*. 1981;155:7–20.
3. Angst F, Schwyzler HK, et al. Measures of adult shoulder function. *Arthritis Care Res*. 2011;63(Suppl 11):S174–88.
4. Barber-Westin SD, Noyes FR, et al. Rigorous statistical reliability, validity, and responsiveness testing of the Cincinnati knee rating system in 350 subjects with uninjured, injured, or anterior cruciate ligament-reconstructed knees. *Am J Sports Med*. 1999;27(4):402–16.
5. Beaton DE, Wright JG, Katz JN, The Upper Extremity Collaborative Group. Development of the QuickDASH: comparison of three item-reduction approaches. *J Bone Joint Surg Am*. 2005;87:1038–46.
6. Bellamy N. WOMAC Osteoarthritis Index user guide. London: University of Western Ontario; 1995. p. 79.
7. Bellamy N. WOMAC Osteoarthritis Index user guide. Version V. Brisbane: CONROD, The University of Queensland; 2002.
8. Bennett PJ, Patterson C, et al. Development and validation of a questionnaire designed to measure foot-health status. *J Am Podiatr Med Assoc*. 1998;88:419–28.
9. Bieri D, Reeve R, et al. The Faces Pain Scale for the self-assessment of the severity of pain experienced by children: development, initial validation and preliminary investigation for ratio scale properties. *Pain*. 1990;41:139–50.
10. Binkley JM, Stratford PW, et al. The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. *Phys Ther*. 1999;79:371–83.
11. Brophy RH, Beauvais RL, et al. Measurement of shoulder activity level. *Clin Orthop Relat Res*. 2005;439:101–8.
12. Brophy RH, Levy B, et al. Shoulder activity level varies by diagnosis. *Knee Surg Sports Traumatol Arthrosc*. 2009;17:1516–21.
13. Budiman-Mak E, Conrad K, et al. Theoretical model and Rasch analysis to develop a revised Foot Function Index. *Foot Ankle Int*. 2006;27(7):519–27.
14. Budiman-Mak E, Conrad KJ, et al. The Foot Function Index: a measure of foot pain and disability. *J Clin Epidemiol*. 1991;44(6):561–70.
15. Busija L, Osborne RH, et al. Magnitude and meaningfulness of change in SF-36 scores in four types of orthopaedic surgery. *Health Qual Life Outcomes*. 2008;6:55.
16. Celik D. Psychometric properties of the Mayo Elbow Performance Score. *Rheumatol Int*. 2015;35(6):1015–20.
17. Cella D, Riley W, et al. Initial Adult Health Item Banks and First Wave Testing of the Patient-Reported Outcomes Measurement Information System (PROMISTM) network: 2005–2008. *J Clin Epidemiol*. 2010;63(11):1179–94.
18. Christensen CP, Althausen PL, Mittleman MA, et al. The nonarthritic hip score: reliable and validated. *Clin Orthop Relat Res*. 2003;406:75–83.
19. Chung KC, Pillsbury MS, Walters MR, Hayward RA. Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. *J Hand Surg Am*. 1998;23(4):575–87.
20. Clarsen B, Myklebust G, et al. Development and validation of a new method for the registration of overuse injuries in sports injury epidemiology: the Oslo Sports Trauma Research Centre (OSTRC) Overuse Injury Questionnaire. *Br J Sports Med*. 2013;47:495–502.
21. Collins NJ, Misra D, Felson DT, Crossley KM, Roos EM. Measures of knee function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Knee Outcome Survey Activities of Daily Living Scale (KOS-ADL), Lysholm Knee Scoring Scale, Oxford Knee Score (OKS), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), Activity Rating Scale (ARS), and Tegner Activity Score (TAS).

- Arthritis Care Res (Hoboken). 2011;63(Suppl 11):S208–28. <https://doi.org/10.1002/acr.20632>.
22. Constant CR, Gerber C, Emery RJ, Sojbjerg JO, Gohlke F, Boileau P. A S186 Angst et al review of the Constant score: modifications and guidelines for its use. *J Shoulder Elb Surg*. 2008;17:355–61.
 23. Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop Relat Res*. 1987;214:160–4. 56.
 24. Cook JJ, Cook EA, et al. Validation of the American College of Foot and Ankle Surgeons Scoring Scales. *J Foot Ankle Surg*. 2011;50(4):420–9.
 25. Cooney WP, Bussey R, Dobyns JH, Linscheid RL. Difficult wrist fractures. Perilunate fracture-dislocations of the wrist. *Clin Orthop Relat Res*. 1987;(214):136–47.
 26. Craig CL, Marshall AL, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc*. 2003;35(8):1381–95.
 27. Crossley KM, Bennell KL, et al. Analysis of outcome measures for persons with patellofemoral pain: which are reliable and valid? *Arch Phys Med Rehabil*. 2004;85(5):815–22.
 28. Dacombe PJ, Amirfeyz R, Davis T. Patient-reported outcome measures for hand and wrist trauma: is there sufficient evidence of reliability, validity, and responsiveness? *Hand (New YorkNY)*. 2016;11(1):11–21.
 29. Daughton DM, Fix AJ, et al. Maximum oxygen consumption and the ADAPT quality-of life scale. *Arch Phys Med Rehabil*. 1982;63:620–2.
 30. Dawson J, Doll H, Boller I, et al. The development and validation of a patient-reported questionnaire to assess outcomes of elbow surgery. *J Bone Joint Surg Br*. 2008;90:466–73.
 31. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg Br*. 1996;78:185–90.
 32. Dawson J, Fitzpatrick R, Carr A. Questionnaire on the perceptions of patients about shoulder surgery. *J Bone Joint Surg Br*. 1996;78:593–600.
 33. Dawson J, Fitzpatrick R, Carr A. The assessment of shoulder instability. The development and validation of a questionnaire. *J Bone Joint Surg Br*. 1999;81:420–6.
 34. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br*. 1998;80:63–9.
 35. Dawson J, Rogers K, Fitzpatrick R, Carr A. The Oxford Shoulder Score revisited. *Arch Orthop Trauma Surg*. 2009;129:119–23.
 36. Dreiser R, Maheu E, Guillou GB, Caspard H, Grouin JM. Validation of an algofunctional index for osteoarthritis of the hand. *Rev Rhum Engl Ed*. 1995;62:43S–53S.
 37. Ferreira-Valente MA, Pais-Ribeiro JL, et al. Validity of four pain intensity rating scales. *Pain*. 2011;152(10):2399–404.
 38. Fix AJ, Daughton DM. Human activity profile professional manual. Odessa: Psychological Assessment Resources, Inc.; 1988.
 39. Glazer DD. Development and preliminary validation of the Injury-Psychological Readiness to Return to Sport (I-PRRS) Scale. *J Athl Train*. 2009;44(2):185–9.
 40. Goldstein CL, Schemitsch E, et al. Comparison of different outcome instruments following foot and ankle trauma. *Foot Ankle Int*. 2010;31(12):1075–80.
 41. Golightly YM, Devellis RF, et al. Psychometric properties of the foot and ankle outcome score in a community-based study of adults with and without osteoarthritis. *Arthritis Care Res*. 2014;66(3):395–403.
 42. Halasi T, Kynsburg A, et al. Development of a new activity score for the evaluation of ankle instability. *Am J Sports Med*. 2004;32:899–908.
 43. Hamamoto Y, Ito H, et al. Cross-cultural adaptation and validation of the Japanese version of the new Knee Society Scoring System for osteoarthritic knee with total knee arthroplasty. *J Orthop Sci*. 2015;20(5):849–53.
 44. Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. *J Bone Joint Surg Am*. 1969;51:737–55.
 45. Hasenstein T, Greene T, et al. A 5-year review of clinical outcome measures published in the Journal of the American Podiatric Medical Association and the Journal of Foot and Ankle Surgery. *J Foot Ankle Surg*. 2017;56(3):519–21.
 46. Hawker GA, Mian S, et al. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short Form McGill Pain Questionnaire (SF MPQ), Chronic Pain Grade Scale (CPGS), Short Form 36 Bodily Pain Scale (SF 36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care Res*. 2011;63:S240–52.
 47. Herdman M, Gudex C, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727–36.
 48. Hernandez-Sanchez S, Hidalgo MD, et al. Responsiveness of the VISA-P scale for patellar tendinopathy in athletes. *Br J Sports Med*. 2014;48(6):453–7.
 49. Hicks CL, von Baeyer CL, et al. The Faces Pain Scale – Revised: toward a common metric in pediatric pain measurement. *Pain*. 2001;93:173–83.
 50. Hunt KJ, Hurwit D. Use of patient-reported outcome measures in foot and ankle research. *J Bone Joint Surg Am*. 2013;95(16):e118(1–9).
 51. Hunt KJ, Lakey E. Patient-reported outcomes in foot and ankle surgery. *Orthop Clin North Am*. 2018;49(2):277–89.

52. American Academy of Orthopaedic Surgeons. Instruments for collection of orthopaedic quality data; 2016.
53. Irrgang JJ, Anderson AF, Boland AL, Harner CD, Kurosaka M, Neyret P, et al. Development and validation of the International Knee Documentation Committee subjective knee form. *Am J Sports Med.* 2001;29:600–13.
54. ISAKOS Scientific Committee, Audigé L, Ayeni OR, Bhandari M, Boyle BW, Briggs KK, Chan K, Chaney-Barclay K, Do HT FM, Fu FH, Goldhahn J, Goldhahn S, Hidaka C, Hoang-Kim A, Karlsson J, Krych AJ, RF LP, Levy BA, Lubowitz JH, Lyman S, Ma Y, Marx RG, Mohtadi N, Marcheggiani Muccioli GM, Nakamura N, Nguyen J, Poehling GG, Roberts LE, Rosenberg N, Shea KP, Sohani ZN, Soudry M, Voineskos S, Zaffagnini S, International Society of Arthroscopy, Knee Surgery and Orthopaedic Sports Medicine. A practical guide to research: design, execution, and publication. *Arthroscopy.* 2011;27(4 Suppl):S1–112.
55. Jenkinson C, Fitzpatrick R, et al. The Nottingham Health Profile: an analysis of its sensitivity in differentiating illness groups. *Soc Sci Med.* 1988;27:1411–4.
56. Johanson NA, Liang MH, et al. American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments. Reliability, validity, and sensitivity to change. *J Bone Joint Surg Am.* 2004;86-a(5):902–9.
57. Kemp JL, Collins NJ, Roos EM, Crossley KM. Psychometric properties of patient-reported outcome measures for hip arthroscopic surgery. *Am J Sports Med.* 2013;41:2065–73.
58. Kennedy CA, Beaton DE, Solway S, McConnell S, Bombardier C. The DASH outcome measure user's manual. 3rd ed. Toronto: Institute for Work & Health; 2011.
59. Kersten P, White PJ, et al. Is the pain visual analogue scale linear and responsive to change? An exploration using Rasch analysis. *PLoS One.* 2014;9(6):e99485.
60. Kirkley A, Griffin S, Alvarez C. The development and evaluation of a disease-specific quality of life measurement tool for rotator cuff disease: the Western Ontario Rotator Cuff Index (WORC). *Clin J Sport Med.* 2003;13:84–92.
61. Kirkley A, Griffin S, et al. Scoring systems for the functional assessment of the shoulder. *Arthroscopy.* 2003;19(10):1109–20.
62. Kirkley A, Griffin S, McLintock H, Ng L. The development and evaluation of a disease-specific quality of life measurement tool for shoulder instability: the Western Ontario Shoulder Instability Index (WOSI). *Am J Sports Med.* 1998;26:764–72.
63. Kirkley A, Werstine R, Ratjek A, Griffin S. Prospective randomized clinical trial comparing the effectiveness of immediate arthroscopic stabilization versus immobilization and rehabilitation in first traumatic anterior dislocations of the shoulder: long-term evaluation. *Arthroscopy.* 2005;21:55–63.
64. Kitaoka HB, Alexander I, et al. Clinical rating systems for the ankle-hindfoot, midfoot, hallux, and lesser toes. *Foot Ankle Int.* 1994;15(7):349–53.
65. Klassbo M, Larsson E, Mannevik E. Hip Disability and Osteoarthritis Outcome Score: an extension of the Western Ontario and McMaster Universities Osteoarthritis Index. *Scand J Rheumatol.* 2003;32:46–51.
66. Kujala UM, Jaakola LH, Koskinen SK, Taimela S, Hurme M, Nelimarkka O. Scoring of patellofemoral disorders. *Arthroscopy.* 1993;9:159–63.
67. Lee JS, Hobden E, et al. Clinically important change in the visual analogue scale after adequate pain control. *Acad Emerg Med.* 2003;10(10):1128–30.
68. Lequesne M. Indices of severity and disease activity for osteoarthritis. *Semin Arthritis Rheum.* 1991;20(Suppl 2):48–54.
69. Lequesne MG, Mery C, Samson M, Gerard P. Indexes of severity for osteoarthritis of the hip and knee: validation—value in comparison with other assessment tests. *Scand J Rheumatol Suppl.* 1987;65:85–9.
70. Lequesne MG. The algofunctional indices for hip and knee osteoarthritis. *J Rheumatol.* 1997;24:779–81.
71. Lippitt SB, Harryman DT, Matsen FA. A practical tool for evaluation of function: the Simple Shoulder Test. In: Matsen FA, Fu FH, Hawkins RJ, editors. *The shoulder: a balance of mobility and stability.* Rosemont: American Academy of Orthopedic Surgeons; 1993. p. 545–59.
72. Lo IKY, Griffin S, Kirkley A. The development and evaluation of a disease-specific quality of life measurement tool for osteoarthritis of the shoulder: The Western Ontario Osteoarthritis of the Shoulder Index (WOOS). *Osteoarthritis Cartilage.* 2001;9:771–8.
73. Longo G, Franceschi F, Loppini M, Maffulli N, Denaro V. Rating systems for evaluation of the elbow. *Br Med Bull.* 2008;87(1):131–61.
74. Lysholm J, Gillquist J. Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. *Am J Sports Med.* 1982;10:150–4.
75. Macdermid J. Update: the Patient-rated Forearm Evaluation Questionnaire is now the Patient-rated Tennis Elbow Evaluation. *J Hand Ther.* 2005;18:407–10.
76. Martin R, Burdett R, et al. Development of the Foot and Ankle Disability Index (FADI). *J Orthop Sports Phys Ther.* 1999;29(1):A32–3.
77. Martin RL, Irrgang JJ, et al. Evidence of validity for the Foot and Ankle Ability Measure (FAAM). *Foot Ankle Int.* 2005;26(11):968–83.
78. Marx RG, Stump TJ, et al. Development and evaluation of an activity rating scale for disorders of the knee. *Am J Sports Med.* 2001;29(2):213–8.
79. Melzack R. The short-form McGill Pain Questionnaire. *Pain.* 1987;30:191–7.
80. Michener LA, McClure PW, Sennett BJ. American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form, patient self-report

- section: reliability, validity, and responsiveness. *J Shoulder Elb Surg.* 2002;11(6):587–94.
81. Mohtadi NGH, Griffin DR, Pedersen ME, et al. The development and validation of a self-administered quality-of-life outcome measure for young, active patients with symptomatic hip disease: the International Hip Outcome Tool (iHOT-33). *Arthroscopy.* 2012;28(5):595–610.e1.
 82. Morrey BF, An KN. Functional evaluation of the elbow. In: Morrey BF, editor. *The elbow and its disorders.* Philadelphia: WB Saunders; 1993.
 83. Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr AJ, et al. The use of the Oxford Hip and Knee Scores. *J Bone Joint Surg Br.* 2007;89:1010–4.
 84. Narin S, Unver B, et al. Cross-cultural adaptation, reliability and validity of the Turkish version of the Hospital for Special Surgery (HSS) Knee Score. *Acta Orthop Traumatol Turc.* 2014;48(3):241–8.
 85. Nilsson AK, Lohmander LS, Klassbo M, Roos EM. Hip Disability and Osteoarthritis Outcome Score (HOOS): validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord.* 2003;4:10.
 86. Nilsson A, Bremander A. Measures of hip function and symptoms: Harris Hip Score (HHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Lequesne Index of Severity for Osteoarthritis of the Hip (LISOH), and American Academy of Orthopedic Surgeons (AAOS) Hip and Knee Questionnaire. *Arthritis Care Res.* 2011;63:S200–7.
 87. Nilsson, et al. The Swedish version of OMAS is a reliable and valid outcome measure for patients with ankle fractures. *BMC Musculoskelet Disord.* 2013;14:109.
 88. Olerud C, Molander H. A scoring scale for symptom evaluation after ankle fracture. *Arch Orthop Trauma Surg.* 1984;103:190–4.
 89. Perruccio AV, Stefan Lohmander L, Canizares M, Tennant A, Hawker GA, Conaghan PG, et al. The development of a short measure of physical function for knee OA KOOS-Physical Function Shortform (KOOS-PS): an OARSI/OMERACT initiative. *Osteoarthritis Cartil.* 2008;16:542–50.
 90. Pinski E, Daniels TR. AOFAS position statement regarding the future of the AOFAS clinical rating systems. *Foot Ankle Int.* 2011;32(9):841–2.
 91. Ramisetty N, Kwon Y, et al. Patient-reported outcome measures for hip preservation surgery: a systematic review of the literature. *J Hip Preserv Surg.* 2015;2(1):15–27.
 92. Richards RR, An KN, Bigliani LU, Friedman RJ, Gartsman GM, Gristina AG, et al. A standardized method for the assessment of shoulder function. *J Shoulder Elb Surg.* 1994;3:347–52.
 93. Richardson J, Sinha K, et al. Modeling the utility of health states with the Assessment of Quality of Life (AQoL) 8D instrument: overview and utility scoring algorithm. Research paper, vol. 63. Melbourne: Centre for Health Economics, Monash University; 2011.
 94. Roach KE, Budiman-Mak E, Songsiridej N, Lertratanakul Y. Development of a Shoulder Pain and Disability Index. *Arthritis Care Res.* 1991;4:143–9.
 95. Robinson JM, Cook JL, et al. The VISA-A questionnaire: a valid and reliable index of the clinical severity of Achilles tendinopathy. *Br J Sports Med.* 2001;35(5):335–41.
 96. Roelofs J, van Breukelen G, et al. Norming of the Tampa Scale for Kinesiophobia across pain diagnoses and various countries. *Pain.* 2011;152:1090–5.
 97. Rompe JD, Overend TJ, et al. Validation of the Patient-rated Tennis Elbow Evaluation Questionnaire. *J Hand Ther.* 2007;20(1):3–10.
 98. Roos EM, Brandsson S, et al. Validation of the foot and ankle outcome score for ankle ligament reconstruction. *Foot Ankle Int.* 2001;22(10):788–94.
 99. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS): development of a self-administered outcome measure. *J Orthop Sports Phys Ther.* 1998;28:88–96.
 100. Rowan K. The development and validation of a multi-dimensional measure of chronic foot pain: the Rowan Foot Pain Assessment Questionnaire (ROFPAQ). *Foot Ankle Int.* 2001;22:795–809.
 101. Scott J, Huskisson EC. Graphic representation of pain. *Pain.* 1976;2(2):175–84.
 102. Scuderi GR, Bourne RB, Noble PC, Benjamin JB, Lonner JH, Scott WN. The New Knee Society Knee Scoring System. *Clin Orthop Relat Res.* 2012;470(1):3–19.
 103. Seeger J, Weinmann S, et al. The Heidelberg Sports Activity Score - a new instrument to evaluate sports activity. *Open Orthopaed J.* 2013;7:25–32.
 104. SooHoo NF, Vyas R, et al. Responsiveness of the foot function index, AOFAS clinical rating systems, and SF-36 after foot and ankle surgery. *Foot Ankle Int.* 2006;27(11):930–4.
 105. Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. *Clin Orthop Relat Res.* 1985;(198):43–9.
 106. Thomas JL, Christensen JC, et al. ACFAS Scoring Scale user guide. *J Foot Ankle Surg.* 2005;44(5):316–35.
 107. Thorborg K, Holmich P, Christensen A, Petersen F, Roos EM. The Copenhagen Hip and Groin Outcome Score (HAGOS): development and validation according to the COSMIN checklist. *Br J Sports Med.* 2011;45:478–91.
 108. van der Linde JA, van Kampen DA, et al. The Oxford Shoulder Instability Score; validation in Dutch and first-time assessment of its smallest detectable change. *J Orthop Surg Res.* 2015;10:146.
 109. Walker N, et al. A preliminary development of the Re-Injury Anxiety Inventory (RIAI). *Phys Ther Sport.* 2010;11(1):23–9.

110. Ware J Jr, Kosinski M, et al. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34(3):220–33.
111. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30(6):473–83.
112. Wendel-Vos GCW, Schuit AJ, et al. Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. *J Clin Epidemiol*. 2003;56:1163e9.
113. Williams GN, Molloy JM, et al. Evaluation of the Sports Ankle Rating System in young, athletic individuals with acute lateral ankle sprains. *Foot Ankle Int*. 2003;24:274–82.
114. Zahiri CA, Schmalzried TP, et al. Assessing activity in joint replacement patients. *J Arthroplast*. 1998;13(8):890–5.
115. Zampelis V, Ornstein E, et al. A simple visual analog scale for pain is as responsive as the WOMAC, the SF-36, and the EQ-5D in measuring outcomes of revision hip arthroplasty. *Acta Orthop*. 2014;85(2):128–32.
116. Zwiers R, Weel H, et al. Large variation in use of patient-reported outcome measures: a survey of 188 foot and ankle surgeons. *Foot Ankle Surg*. 2017. pii: S1268-7731(17)30050-4.