American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 72, No. S10, October 2020, pp 200–218
DOI 10.1002/acr.24231

Measures of Hip Function and Symptoms

Jack W. Weick, Jace Bullard, Jordan H. Green, and Joel J. Gagnier

INTRODUCTION

Accurate measurements of hip outcomes and function are vital in the study of hip pathology. Additionally, outcome measures can also be used in the clinical realm to track progress of treatment. Several outcome measures have been developed in an effort to accurately assess hip symptoms and function. A 2011 review of the most commonly used hip outcome scores was previously published in this journal (1). Since then, a number of additional studies focusing on these outcome scores have been published to further refine our knowledge and potential utility of hip outcome scores. The purpose of this review is to provide updated information regarding the most commonly used hip outcome scores, which include the Harris Hip Score (HHS), the Hip Disability and Osteoarthritis Outcome Score (HOOS), the Patient-Reported Outcomes Measurement Information System (PROMIS), the Oxford Hip Score (OHS), the Lequesne Index of Severity for Osteoarthritis of the Hip (LISOH), the American Academy of Orthopaedic Surgeons (AAOS) Hip and Knee Questionnaire, and the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). Table 1 provides information on practical applications of these scores. Table 2 provides a summary of the psychometric properties of the included hip outcome scores.

HARRIS HIP SCORE

Description

Purpose. The HHS was initially developed for the assessment of the results of hip surgery, specifically mold arthroplasties for posttraumatic osteoarthritis of the hip, in 1969 (2). It is now intended to evaluate various hip disabilities and methods of treatment in the adult population.

Content. The original HHS covers pain, function, absence of deformity, and range of motion. The pain domain measures pain severity and its effect on activities and need for pain medication. The function domain assesses daily activities (stair use, using public transportation, sitting, and managing shoes and socks) and gait. Deformity evaluates hip flexion, adduction, internal rotation, and extremity length discrepancy. Range of motion measures hip flexion, abduction, external and internal rotation, and adduction.

A modified version is also now available that only assesses the pain and function components of the HHS (3,4).

Number of items. In the original HHS, there are 10 items. In the modified HHS, there are eight items.

Response options/scale. The original HHS has a maximum of 100 points (best possible outcome) covering pain (one item, 0-44 points), function (seven items, 0-47 points), absence of deformity (one item, 4 points), and range of motion (two items, 5 points). The modified HHS only includes pain and function, so it has a maximum of 91 points; however, the score is multiplied by 1.1 to get a maximum of 100.

Recall period for items. Not described.

Cost to use. Free.

How to obtain. The HHS is available in the original article (2) and at http://www.orthopaedicscore.com/. The modified HHS can be obtained by only completing the pain and function components and then multiplying the score by 1.1.

Practical application

Method of administration. The HHS is a clinician-based outcome measure administered by a qualified health care professional, such as a physician or a physical therapist. The modified HHS is a self-administered instrument.

Scoring. Each item has a unique numerical scale that corresponds to descriptive response options. The number of response options as well as the number of points assigned varies by each item. In the original HHS, the range of motion item consists of six motions graded. Each range of motion factor is assigned an

Jack W. Weick, MD, Jace Bullard, MD, Jordan H. Green, MD, Joel J. Gagnier, ND, MSc, PhD: University of Michigan, Ann Arbor.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Joel J. Gagnier, ND, MSc, PhD, MedSport Sports Medicine at Domino's Farms, 24 Frank Lloyd Wright Drive, Ann Arbor, MI 48105. E-mail: jgagnier@med.umich.edu.

Submitted for publication March 5, 2020; accepted in revised form April 21, 2020.

index factor and a maximum possible value. These points are added and multiplied by 0.05 to receive the total points for range of motion. The total score is calculated by summing the scores for the four domains. The modified HHS is calculated by summing the pain and function components of the scale and multiplying by 1.1.

Score interpretation. Both the HHS and modified HHS have a maximum of 100 points. The higher the HHS, the less dysfunction. Scores below 70 are typically considered a poor result. A score of 70 to 80 is considered fair, 80 to 90 is considered good, and 90 to 100 is considered excellent (2). Pain and function are the major score contributors, with 44 points possible for pain and 47 points possible for function (14 for activities of daily living and 33 for gait). Range of motion has a maximum of 5 points, and deformity has a maximum of 4 points.

Respondent time to complete. The HHS takes approximately 5 minutes to complete.

Administrative burden. No formal training is necessary to administer the HHS.

Translations/adaptations. The HHS has been used internationally (United States, Canada, Sweden, Europe, etc), but there are no validated versions in other languages.

Psychometric information

Floor and ceiling effects. Unacceptable ceiling effects for the HHS were reported in 31 of 59 studies (5). Pooled data across studies (n = 6667 patients) suggested ceiling effects of 20% (95% confidence interval 18-22). The modified HHS had reported ceiling effects of 27.5% (6). In a population of 294 patients who underwent periacetabular osteotomy, the modified HHS had reported ceiling effects of 1% (7).

Reliability. For the original HHS, the test-retest reliability for the total score was excellent for physicians (r = 0.94) and physiotherapists (r = 0.95). Both physicians and physiotherapists had excellent test-retest reliability for pain and function (r = 0.93-0.98) (8). The interrater correlations were good to excellent in two previous studies (0.74-1.0) (8,9).

The modified HHS was shown to have excellent reliability, with Cronbach's α of 0.95 and high intraclass correlation (0.91-0.95) (10).

Validity. The content validity of the HHS has been tested by comparisons with the WOMAC and the 36-item Short Form Health Survey (SF-36). The HHS has demonstrated no major differences when compared with these scores (8). In assessment of construct validity, the pain and function domains in the HHS have been shown to correlate with similar domains in the WOMAC, the Nottingham Health Profile, and the SF-36 (8,11). The correlation between the HHS and SF-36 is notably strong in the physical domains and weak in the mental domains.

The modified HHS has been compared with the 12-item Short Form Health Survey (SF-12) physical and mental subscale for validity and has been shown to have poor correlation with the mental subscale but strong correlation with the physical subscale (7).

Responsiveness. In a study of 335 total hip arthroplasties (THAs), the HHS was found to be responsive to pain and function at the 6-month postoperative follow-up but weak at the 2-year follow-up (12). In a study of 293 periacetabular osteotomies, the modified HHS was found to be adequately responsive (7).

Minimally important differences. In a study assessing the HHS in femoral neck fractures 4 and 12 months postoperatively, the standardized response mean (SRM) was 0.75, which was the best in ability to detect change when compared with the Barthel Index and the EQ-5D (13). The modified HHS was shown to have the lowest minimal clinically important difference (MCID) when compared with the WOMAC and HOOS (7).

Generalizability. Both the HHS and the modified HHS are generalizable to adult patients with hip disabilities or undergoing hip surgeries.

Use in clinical trials. Both the HHS and the modified HHS have been used extensively in clinical trials involving hip pathology and treatment of different hip conditions. The HHS has been used in trials that range from assessment of treatment choice for femoral neck and intertrochanteric femur fractures to assessment of effectiveness of different injections for osteoarthritis of the hip (14–17). The HHS is also a popular tool in studies relating to THAs given the focus of the measure on pain and impaired physical function, which are the main indications for a THA (18–20). Similarly, the modified HHS is widely used in clinical trials relating to hip pathology, including injections, management of hip fractures, hip arthroscopy, and THAs (21–25). The modified HHS has the benefit in clinical trials, over the HHS, of being self-administered and not requiring a clinician for assessment.

Critical appraisal of overall value to the rheumatology community

Strengths. As discussed previously, the HHS and modified HHS are both widely used in evaluating outcomes of THA given the focus on evaluation of pain and physical function (26). Additionally, the HHS and modified HHS have been shown to be appropriate for use in measurement of outcomes of intervention for treatment of femoral neck fractures, hip arthroscopy, and conservative management of hip osteoarthritis (7,13,27).

Caveats and cautions. The main criticism of the HHS is the ceiling effects, which limit its validity (5).

Clinical usability. The psychometric evaluation does not support interpretation of scores to make decisions for individuals.

Research usability. As long as the researcher is aware of the ceiling effects, both the HHS and modified HHS can be used for clinical outcome studies for a variety of hip pathologies.

HIP DISABILITY AND OSTEOARTHRITIS OUTCOME SCORE

Description

Purpose. The HOOS was developed as an adaptation of the Knee Injury and Osteoarthritis Outcome Score (KOOS), which itself is an adaptation of the WOMAC. It was created as an instrument of assessment of adult patients' opinions about their hip and hip disability. The HOOS was originally developed in 2003 and has been validated in two different versions (28).

Several variations of the HOOS have also been developed: the HOOS–Joint Replacement Short Form (HOOS-JR), the HOOS–Physical Function Short Form (HOOS-PS) (29,30), and the 12-item HOOS (HOOS-12). The HOOS-JR is a six-item instrument developed specifically for patients undergoing THA that measures the domains of pain and activities of daily living (ADLs). The HOOS-PS is a five-item measure of physical function designed to elicit patients' opinions about difficulties experienced because of their hip problems. The HOOS-12 retains the three subscores (pain, function/ADL, and quality of life [QOL]).

Content. The HOOS consists of five subscales: pain, other symptoms, function in ADLs, function in sport and recreation, and hip-related QOL. The HOOS-JR measures only the domains of pain and ADLs. The HOOS-PS measures physical function.

Number of items. The original HOOS contains 40 items total: 10 for pain, 5 for other symptoms, 17 for function in ADLs, 4 for function in sport and recreation, and 4 for hip-related QOL. The HOOS-JR contains six items total: two for pain and four for function. The HOOS-PS contains five items total, all relating to physical function. The HOOS-12 has 12 items: 4 for pain, 4 for function, and 4 for QOL.

Response options/scale. In the HOOS, HOOS-JR, and HOOS-PS, standardized answer options are given in five Likert boxes. Each question is scored 0 to 4. Scores are summarized and transformed onto a 0 to 100 scale. A score of 0 indicates extreme problems, and a score of 100 indicates no problems.

Recall period for items. The last week is taken into consideration for the questions.

Cost to use. Free.

How to obtain. The HOOS, HOOS-PS, and HOOS-12 can be obtained at www.koos.nu. The HOOS-JR can be obtained at www.aaos.org/uploadedFiles/HOOS-JR-2016.pdf.

Practical application

Method of administration. The questionnaire for all forms of the HOOS is patient reported.

Scoring. The manual scoring sheet is available from the above websites. Included are instructions for handling missing values. Computer scoring can increase clinical usefulness.

Score interpretation. Scores can range from 0 to 100, and each subscale can have scores from 0 to 100, in which 0 indicates extreme problems and 100 indicates no problems. The HOOS scores within each subscale (pain, symptoms, ADLs, sport and recreation, and QOL) can be plotted for comparison of pre- and postintervention comparison visualization.

Respondent time to complete. The original HOOS takes 10 to 15 minutes for a subject to complete. The HOOS-JR, HOOS-PS, and HOOS-12 all take less than 5 minutes to complete.

Administrative burden. There is minimal administrative burden. Hand scoring of the HOOS can take 10 to 15 minutes without additional training. Computerized scoring can automate this process for instantaneous results without further administrative need.

Translations/adaptations. The HOOS is available in 25 total languages, including English, French, German, Dutch, Italian, and Spanish, among others (available on www.koos.nu). A number of adaptations of the HOOS exist, including the KOOS, the KOOS for children, the Foot and Ankle Outcome Score, the Rheumatoid and Arthritis Outcome Score, the Hip and Groin Outcome Score, and the Neck Outcome Score.

Psychometric information

Floor and ceiling effects. Floor and ceiling effects have been reported in different subscales within the HOOS. Floor effects have been reported to range from 4.1% to 17.8% in the sport and recreation subscale in subjects eligible for THA and patients with hip osteoarthritis (31–33). Ceiling effects have been

reported in THA in all subscales of the HOOS (31). No floor effects have been reported in the hip arthroscopy population; however, ceiling effects were reported 12 and 24 months postoperatively in the ADL and sport and recreation subscales (34).

The HOOS-JR has been shown to have low floor effects of 0.6% to 1.9% but ceiling effects of up to 37% to 46% after THA (29). Floor and ceiling effects of the HOOS-PS have only been studied in the French translation of the HOOS-PS, but no floor or ceiling effects were observed (33).

Reliability. The reliability of the HOOS has been examined in patients treated conservatively for hip osteoarthritis, in patients treated with THA, and in patient undergoing hip arthroscopy (28,31–34). In patients undergoing THA, the internal consistency ranged from 0.82 to 0.98 (Cronbach's α coefficient), with the ADL subscale having the highest consistency of 0.94 to 0.98 (28,32,33). Test-retest reproducibility has been shown to be high in the THA population, with an intraclass correlation coefficient of 0.75 to 0.97 (28,32,33). In the hip arthroscopy population, test-retest reliability was excellent, with intraclass correlation coefficients ranging from 0.91 to 0.97 (34).

The HOOS-JR was shown to have an acceptable internal consistency, as measured by the Person Separation Index, of 0.86 (29). Test-retest reliability was not reassessed in the HOOS-JR. The HOOS-PS was shown to have an intraclass correlation coefficient of 0.86 (33).

Validity. HOOS content validity was performed by asking patients to rate item importance (28,31). Construct validity has been confirmed in hip osteoarthritis, THA, and hip arthroscopy studies by comparison with the SF-36, OHS, LISOH, and visual analog scale (VAS) for pain (31–34). HOOS-JR and HOOS-12 construct validity was confirmed, with high correlations with the pain and ADL domains of the HOOS as well as the pain and function domains of the WOMAC and moderate correlations with the other domains of the HOOS and WOMAC (29).

Responsiveness. All domains of the HOOS were found to be responsive in hip arthroscopy at the 9- and 12-month postoperative follow-up (34). In a study of patients who underwent a periacetabular osteotomy, the HOOS was found to be responsive at the 12-month follow-up (7). The HOOS-JR and HOOS-12 were found to have very high responsiveness up to 2 years postoperatively (29). The HOOS-PS was also found to have good responsiveness but was only assessed 1 month postoperatively (33).

Minimally important differences. The smallest detectable difference of the HOOS ranged from 9.6 in the ADL domain to 16.2 in the QOL domain (33). In hip arthroscopy, the minimal detectable change ranged from 9 in the ADL domain to 19 in the QOL domain (34).

Generalizability. The HOOS is generalizable to adult populations with hip ailments or undergoing hip procedures. The HOOS-JR was developed for patients undergoing THA.

Use in clinical trials. The HOOS has been used extensively in clinical trials relating to hip osteoarthritis and hip pathology. The HOOS is one of the more common patient-reported measures of hip function and postoperative outcomes. It has been used extensively in THA clinical studies (35–37), hip dysplasia surgery (38,39), hip arthroscopy (40), and even in development of pain management protocols after hip surgery (41). The HOOS-JR has only existed since 2016 and is intended mainly for THA, but it has been used in recent trials related to THA outcomes (42,43). The HOOS-PS has also been used in clinical trials relating to hip surgery, but it is used far less frequently than the original HOOS (18,44,45).

Critical appraisal of overall value to the rheumatology community

Strengths. The HOOS can be used for younger and more active patients given the subscales and individual domains. The HOOS has shown favorable qualities in reviews of psychometric properties in assessment of patients after THA, hip arthroscopy, and periacetabular osteotomy and patients with hip osteoarthritis not undergoing surgery (7,34,46,47).

Caveats and cautions. The HOOS is overall well received; however, it does have ceiling effects in all domains, which are also present in the HOOS-JR (29,34).

Clinical usability. The HOOS can be used to observe patients with hip conditions over time in clinics. The HOOS-JR can be used in clinics that perform THA for monitoring patients.

Research usability. The HOOS, HOOS-JR, HOOS-12, and HOOS-PS are all usable in the research setting for hip conditions.

PATIENT-REPORTED OUTCOMES MEASUREMENT INFORMATION SYSTEM

Description

Purpose. PROMIS was developed in 2004 by the National Institutes of Health for use in clinical care and medical research (48). This system was intended to be used across multiple specialties and was intended to be publicly available, precise, and flexible.

Content. The PROMIS adult profile covers three domains: physical health, mental health, and social health. Within each of these domains are more specific profile domains. Profile domains under physical health include fatigue, pain intensity, pain interfer-

ence, physical function, and sleep disturbance. Profile domains under mental health include anxiety and depression. Within orthopedics, the PROMIS physical function, pain interference, and depression domains have been most commonly used (49–53).

Number of items. There are multiple options of measures within PROMIS, which can vary in the number of items. The short forms are a fixed set of 4 to 10 items for one individual domain. Computer adaptive tests (CATs) provide items selected dynamically from an item bank based on the subject's previous answers. The CATs are usually 4 to 12 items in length. Profiles are fixed collections of short forms from seven different domains. Several versions of PROMIS profiles exist, ranging from four to eight questions per domain.

Response options/scale. PROMIS uses a T-score metric; 50 is designed to be the mean of the population, and 10 is 1 SD of the population. Therefore, a score of 60 would indicate that the subject is 1 SD above the mean for the population within the measured domain. Response options of the questions depend on the category of the question. Pain is typically measured on a 0- to 10-point scale. Frequency-, duration-, intensity-, and capability-related questions typically have a scale of five variable options.

Recall period for items. Most PROMIS items use a 7-day recall period. However, typically within the physical function domain, the question is assessed in the present tense.

Cost to use. Free.

How to obtain. The different PROMIS measures can be obtained from the HealthMeasures website at www.healthmeasures.net/explore-measurement-systems/promis.

Practical application

Method of administration. PROMIS scales are administered via paper or computer-based forms.

Scoring. Scoring manuals are available on the HealthMeasures website to allow for conversion of the score to a T-score. PROMIS measures are scored on a T-score metric; 50 is designed to be the mean of the population, and 10 is 1 SD of the population. Therefore, a score of 60 would indicate that the subject is 1 SD above the mean for the population within the measured domain.

Score interpretation. A score of 50 indicates that the subject measures at the mean of the population in that particular metric; 1 SD is 10 points. Therefore, a score of 60 would indicate that the subject is 1 SD above the mean for the population within the measured domain.

Respondent time to complete. PROMIS domains relating to hip pathology typically take less than 5 to 10 minutes for the subject to complete.

Administrative burden. There is minimal overall administrative burden. However, the distributor must have access to the scoring manuals for interpretation of the subjects' responses.

Translations/adaptations. The PROMIS adult domains are available in more than 45 different languages, including English, Spanish, French, German, Dutch, and Italian.

Psychometric information

Floor and ceiling effects. There has been little overall study on psychometric properties of the PROMIS score in relation to hip pathology. In a study of patients who underwent periacetabular osteotomy for symptomatic hip dysplasia, PROMIS showed no floor or ceiling effects both pre- and postoperatively (39). When studied in patients who underwent THA for symptomatic hip osteoarthritis, the PROMIS depression domain showed floor effects up to 20% preoperatively and 30% to 45% postoperatively (54). In the same study, the PROMIS pain interference domain showed no floor effects preoperatively but showed floor effects of 21% to 26% at 1 year postoperatively. No ceiling effects were observed in the depression, pain interference, or physical function scores pre- or postoperatively.

Reliability. Few studies have assessed the reliability of PROMIS instruments related to treatment of hip pathology. One study examining the psychometric properties of one CAT of the PROMIS (the lower-extremity CAT) compared with the modified HHS and the Hip Outcome Score showed excellent reliability, with a Cronbach's α score of 1.00 (6).

Validity. To our knowledge, content validity of PROMIS instruments has not been evaluated specifically for hip pathology. Group validity has been assessed previously in patients with a general diagnosis of osteoarthritis but has not been assessed specifically for hip osteoarthritis (55).

Responsiveness. Responsiveness of PROMIS instruments for hip pathology has been evaluated in a study of patients who underwent THA. The pain interference and physical function domains were found to have high responsiveness postoperatively in patients with hip osteoarthritis after THA (54).

Minimally important differences. The MCID has typically been defined as 5, half of the normalized SD of reported PROMIS scores (SD 10). In a study of patients who underwent THA, 66% to 78% of patients had an MCID postoperatively in the

pain interference domain, 61% to 75% of patients had an MCID in the physical function domain, and 44% to 46% of patients had an MCID in the depression domain (54).

Generalizability. PROMIS scores are generalizable to adult populations with hip ailments or undergoing hip procedures. There are separate PROMIS instruments for pediatric patients.

Use in clinical trials. Compared with other measures of hip function and symptoms, PROMIS instruments have been used relatively less frequently in clinical trials. However, more recent studies have begun to use PROMIS instruments in assessment of outcomes related to surgical intervention for different hip pathologies. Clinical studies on outcomes after THA have used PROMIS scores for assessment of both functional outcomes and pain improvement postoperatively (56,57). PROMIS instruments have been used in clinical trials assessing baseline disability and functional performance in patients with femoroacetabular impingement (FAI) (58,59). Lastly, PROMIS scores were used in one clinical trial studying postoperative outcomes of correction of mildly dysplastic hips with periacetabular osteotomy (60).

Critical appraisal of overall value to the rheumatology community

Strengths. PROMIS instruments have the benefit of adaptability to patient responses, which enables the questionnaire to be adaptable to individual patients. Additionally, PROMIS instruments are becoming more widely used in all fields of health care. Therefore, if used in hip pathology, it has the potential added benefit of easier interpretation by health care practitioners or researchers who may not be as familiar with other hip-specific outcome measures.

Caveats and cautions. There is a relative paucity of literature involving psychometric properties of PROMIS instruments in patients with hip pathology. Further investigation should be performed in this area. Additionally, high floor effects have been reported in a previous THA study, as mentioned previously (54).

Clinical usability. Although PROMIS instruments have not been extensively used in assessment of outcomes in patients with hip pathology, they can be used to follow patients both pre- and postoperatively.

Research usability. Further evaluation of the psychometric properties of PROMIS instruments, specifically for hip pathology, would be beneficial; however, PROMIS instruments have been used in several clinical trials involving THA, FAI, and hip dysplasia (56–60). Investigators should take note of the previously reported high floor effects in THA populations (54).

OXFORD HIP SCORE

Description

Purpose. The purpose of the OHS is to assess outcome after THA by measuring a patient's perceptions after surgery. It was originally described in 1996 and was updated in 2007 (61,62).

Content. The OHS assesses pain (six items) and function (six items) of the hip in relation to daily activities, such as walking, dressing, slipping, etc.

Number of items. The OHS has 12 items with 5 categories of response; there are no subscales.

Response options/scale. The original scoring system, as described by Dawson et al (61) in 1996, ranged from 1 to 5 (best to worst), with a total score of 12 to 60 (least difficulties to most difficulties). This scoring system was modified in 2007 by Murray et al (62) to new item ranges of 0 to 4 (worse to best), with total scores ranging from 0 to 48 (most difficulties to least difficulties) (62,63).

Recall period for items. During the past 4 weeks.

Examples of use. This patient-reported outcome measure (PROM) has been used in several countries in both clinical and research settings. It has been validated and used in both primary and revision hip replacements (64–75).

How to obtain. The OHS questionnaire is free to use and is available online at http://www.orthopaedicscore.com/scorepages/oxford_hip_score.html. Further information on the OHS and all the other Oxford orthopedic scores can be found at https://phi.uhce.ox.ac.uk/ox_scores.php.

Practical application

Method of administration. The orthopedic hip score questionnaire can be self-administered or completed over the phone (61,62,76).

Scoring. As previously stated, each item (12 items) contains five possible responses. According to the updated scoring system, these five responses are scored from 0 to 4 (worse to best), resulting in a possible overall score range of 0 to 48 (most difficulties to least difficulties). The maximum of two missing values can be accepted and replaced by mean values. Overall scores should not be calculated if more than two items are left unanswered. If a patient marks multiple responses for one item, the worst response should be used for calculation of scores (62).

Score interpretation. Categories for the OHS, based on data from the HHS and with a score range of 0 to 48, have suggested cutoff scores of greater than 41 as excellent, 34 to 41 as good, 27 to 33 as fair, and less than 27 as poor (77). More recent research has shown that a postoperative OHS of greater than 37.5 is associated with a successful outcome (78–80). Additionally, clinicians have attempted to use the OHS as a method of screening out patients who do not require total hip replacement. Neufeld and Masri (81) found that an OHS of 34 or higher was a good predictor of successful nonoperative management of hip arthritis. With use of the classification system described by Kalairajah et al (77), the OHS at 6 months is a useful predictor of early revision THA. A poor score was associated with a revision risk of 7.6%, compared with a revision risk of 0.7% in patients with good/excellent scores (67,77). Lastly, normative values were established in 2015 and published by Hamilton et al (79) for patients undergoing THA. They were able to establish normative values for male and female patients in four age categories (younger than 60, 60–70, 70–80, and older than 80) both preoperatively and 12 months after THA (79).

Respondent burden. The OHS takes between 2 and 15 minutes to complete. Based on patient interviews, issues have been raised regarding item clarity and double-barreled questions (82,83).

Administrative burden. The burden of administering the OHS is minimal. Because the OHS is a patient-reported questionnaire, the time to score is short and involves only elementary arithmetic. No specific training is required to score the OHS.

Translations/adaptations. The OHS has been widely used in many countries. It has been translated and validated in Japanese, Dutch, German, Turkish, Spanish, Mandarin, Italian, and Danish and with the use of an on-site translator (66,80–89). It has also been translated into French, Iranian, and Korean without the supporting validation studies (84–86). The orthopedic Oxford scores also include similar questionnaires for assessing outcome after knee replacement, shoulder replacement, elbow replacement, and ankle replacement as well as questionnaires for assessing shoulder instability.

Psychometric information

Method of development. Questions were made based on patient interviews in which patients with hip arthritis were asked to report their experiences and frustrations. Patients were involved in content validity of the questionnaire (61). The OHS underwent item-response theory testing in 2004 by Fitzpatrick et al (87), and there was an overall good item fit of the data to the Rasch model.

Acceptability. In a 2000 study, Fitzpatrick et al (93) showed that 90% of patients filled out the questionnaire to completion. In general, older patients and patients with more severe medical problems were less likely to complete the questionnaire compared with younger and healthier patients. In their study, the most problems were recorded for the item regarding distance walked before severe pain (88,89). In one study, up to 10% of English-speaking Americans misinterpreted this item. The authors hypothesized that the use of "not at all" as an answer choice implied that the patient never had pain, so this confusion might lead to an overall underestimation of their hip function. Like other hip PROMs, the OHS is subject to statistical ceiling effects (approximately 13.5%); however, very low levels of statistical floor effects are observed with the OHS (90,91).

Reliability. Several studies have investigated the internal consistency of the OHS pre- and postoperatively. The studies have shown the internal consistency of the OHS to be high (range 0.84-0.93) preoperatively and at 3, 6, 12, and 24 months postoperatively (61,92,93). Reproducibility, as measured by the coefficient of repeatability or interclass correlations by using the Bland-Altman method, has also been studied extensively and shown to be consistently strong (61,92,94).

Validity. During the development of the OHS, patients were asked to comment on and include hip-related problems not addressed by the draft questionnaire for content validity (61). High correlations (range 0.67-0.85) have been described when comparing the OHS with other PROMs of pain and function related to hip pathology (61,70,77,90,95,96).

Ability to detect change. The OHS has favorable responsiveness when compared with generic measures, such as the SF-36 and EQ-5D, and disease-specific measures, including the WOMAC and Arthritis Impact Measurement Scales. Effect size varied from 2.1 to 3.1 at 6 to 24 months after THA and was 1.84 after revision THA (61,70,90–93,96,97). In 2007, Murray et al (62) estimated the MCID to be between 3 and 5 points after joint replacement. Further study in 2015 by Beard et al (80) found the minimal detectable change to be 5 points after THA.

Critical appraisal of overall value to the rheumatology community

Strengths. The OHS assesses pain and functional outcomes in patients undergoing THA. It has been shown to provide good psychometric properties and has been reported to be a useful predictor of early revision THA.

Caveats and cautions. As previously discussed, the OHS has some questions that may be misinterpreted, therefore leading to an underestimation of hip pain and function. Additionally, concerns have been raised regarding the lack of items concern-

ing activities requiring a large angle of hip flexion. Lastly, it has been shown that 10% of patients whose first language is English may misinterpret one of the items.

Clinical usability. The questionnaire is quick, easy to use, free, and self-administered. Therefore, clinical usability is high. With that said, a single administration will not provide useful information on an individual; however, with repeated administrations, useful information can be gleaned.

Research usability. Because of its ease of use and high response rate, the OHS is one of the preferred PROMs for large studies on long-term hip replacement outcomes (65).

LEQUESNE INDEX OF SEVERITY FOR OSTEOARTHRITIS OF THE HIP

Description

Purpose. The LISOH was developed in France in the early 1980s to evaluate the severity of hip osteoarthritis in drug trials in an adult French population and the long-term treatment effects for hip osteoarthritis as well as to help in decision-making regarding the need for hip replacement (98). It was modified in 1997 and became known to some as the Lequesne Algofunctional Index.

Content. The LISOH is an index that covers osteoarthritis-specific symptoms and physical function disability. It is a composite measure of aggregating symptoms and function (which are not graded separately) in which pain is analyzed by five items, maximum distance walked by two items, and ADLs by four items (98). This instrument is available in interview-based and self-administered versions and in modified versions with changed scoring and wording (98–101).

Number of items. There are 11 items.

Response options/scale. The score ranges from 0 (no pain or disability) to 24 (maximum pain or disability), and the measure is scored as a sum of all the items (98).

Recall period for items. Not specified.

Examples of use. The LISOH has been used in both clinical and research settings since its development in the 1980s. Clinically, it has been used to assess the severity of hip osteoarthritis and to help with indications for THA (101–103). In the research realm, the LISOH has been used to determine the effectiveness of pharmacologic interventions on hip osteoarthritis and to assess the long-term impact of post-THA rehabilitation (98,104)

How to obtain. The LISOH is free to use and can be accessed at https://oarsi.org/sites/default/files/docs/2013/lequesne_eng_ndex.pdf.

Practical application

Method of administration. The LISOH can be self-administered, interviewer administered, or completed by a clinician during a clinical assessment.

Scoring. The original scoring consists of score ranges from 0 to 8 for each part of the LISOH questionnaire (pain/discomfort, maximum distance walked, and ADL), resulting in a total score range of 0 to 24. A modification in 1991 added a question regarding sexual activity to be included when appropriate, resulting in a total score range of 0 to 28 (98,99).

Score interpretation. With the original scoring consisting of a total score range of 0 to 24 points, where 0 = no handicap, 1 to 4 = mild handicap, 5 to 7 = moderate handicap, 8 to 10 = severe handicap, 11 to 13 = very severe handicap, and 14 and higher = extremely severe handicap. A score greater than 11 to 12 points has been suggested to indicate need for THA (103). The questions are suggested to score disabilities connected with a single hip. There are no indications of how to score in the case of bilateral hip osteoarthritis, which complicates interpretation of the LISOH in those patients (102).

Respondent burden. The LISOH questionnaire takes less than 5 minutes to complete (102,105,106).

Administrative burden. Although scoring of the LISOH questionnaire takes only a few minutes, some training may be required for use of the questionnaire in an interview-based environment to achieve interobserver reproducibility (98,105).

Translations/adaptations. The LISOH has been translated and validated for hip osteoarthritis in English, French, German, Turkish, Korean, Spanish, Greek, Persian, and Portuguese (101,102,107–114).

Psychometric information

Method of development. The LISOH was developed in France by hip specialists in the early 1980s through patient interviews in an adult French population.

Acceptability. Several studies have assessed the LISOH questionnaire using Rasch analysis, and unfavorable results have caused some to question the psychometric properties of the questionnaire. Furthermore, in direct comparison with the WOMAC, the LISOH exhibits worse internal consistency, reliability, and construct validity in multiple patient populations (107,110,111,115). Other issues raised regarding the LISOH questionnaire include the clarity of the items, with one study determining that 2 of 10 patients in a French population required additional explanation to fill out the questionnaire and also noting a poor item-response rate and another study noting an item-response rate of approximately 71% (102,105).

Reliability. In general, internal consistency has been found to be satisfactory for the LISOH composite score (Cronbach's α 0.83-0.84) (101,102,108). With that said, the internal consistency has been shown to be lower for the pain/discomfort part of the LISOH in comparison with the function part (Cronbach's $\alpha = 0.63$ vs. 0.84, respectively) (101). Satisfactory test-retest reliability has been shown for the composite score, with interclass correlation coefficients ranging from 0.51 to 0.96 (101,107,111). Regarding interrater reliability, the interview-based questionnaire had a mean deviation of 0.55 points when rated by two observers (98).

Validity. Construct validity and convergent validity of the LISOH questionnaire have been shown to be inferior to those of other PROMs (47,101,102,107,111).

Ability to detect change. The MCID of the LISOH remains to be elucidated. When using the LISOH to assess the long-term impact of active drug treatment for hip osteoarthritis, an effect size of 1.3 to 1.8 was observed (100).

Critical appraisal of overall value to the rheumatology community

Strengths. The LISOH is quick, free, and easy to access.

Caveats and cautions. Problems have been raised regarding the validity of the LISOH and its utility as a single measure of outcomes after THA.

Clinical usability. Given its poor validity and reliability relative to other patient-reported outcomes, it is not recommended to use the LISOH as the sole outcome measure on an individual patient level.

Research usability. Given its ease of use, availability, and relatively high interrater/interobserver reliability, the LISOH questionnaire may have some utility in the study of THA outcomes in large patient populations.

AMERICAN ACADEMY OF ORTHOPAEDIC SURGEONS HIP AND KNEE QUESTIONNAIRE

Description

Purpose. Throughout the 1990s and early 2000s, the AAOS created a series of questionnaires designed to measure and analyze musculoskeletal outcomes. These assessments covered all body regions in adults and children. The Hip and Knee Questionnaire was a specific version of the more general Lower Limb Questionnaire and was published in 2004. In combination with the SF-36, the AAOS Hip and Knee Questionnaire was effective at assessing hip and knee conditions and the effects of treatment (116). The survey was previously available through the AAOS website but has now been replaced by the HOOS and KOOS surveys.

Content. The questionnaire is identical to the AAOS Lower Limb Questionnaire with the attribution of pain to the hip or knee. It asks respondents to answer questions regarding hip or knee stiffness, swelling, and function. Specifically, the functional questions assess the respondent's pain while walking on flat surfaces, going up or down stairs, and lying in bed at night. The final two questions assess the ability to get around and the level of difficulty with taking on and off socks.

Number of items. The questionnaire contains seven items. The worse hip is given preference in the case of bilateral symptoms.

Response options/scale. For questions regarding stiffness and swelling, respondents choose from five possible options on a Likert scale from "not at all" to "extremely." For questions regarding pain during functional activity, there are seven response options on a Likert scale from "not painful" to "could not do" as well as a "could not do for other reasons" option. There are seven response options for ability to get around and six response options for difficulty with taking on and off socks.

Recall period for items. One week.

Cost to use. Free.

How to obtain. Previously available at aaos.org.

Practical application

Method of administration. Self-administered questionnaire.

Scoring. Standardized and normative scores may be calculated. Unanswered items are not to be included in calculation of mean scores. In addition, if more than half of the items are missing, a score cannot be calculated. This includes items marked

"could not do for other reasons," which is considered equal to a missing response. Instructions on how to calculate standardized and normative scores are included in a worksheet with the survey.

Score interpretation. Standardized scores are calculated from 0 to 100, 0 indicating most disability and 100 indicating least disability. Normative scores are calculated to a mean population score of 50, with higher scores indicating higher function (117).

Respondent time to complete. Two to three minutes.

Administrative burden. The questionnaire can be scored very quickly with a scoring sheet and within 15 minutes if scored by hand.

Translations/adaptations. The Hip and Knee Questionnaire is an adaptation of the Lower Limb Questionnaire. There are also sports/knee and foot and ankle adaptations (116).

Psychometric information

Floor and ceiling effects. Not measured (116).

Reliability. The Hip and Knee Questionnaire has good internal consistency, with Cronbach's α coefficient of 0.8 calculated from 43 patients with a hip or knee complaint. For test-retest reliability, Pearson's correlation coefficient was calculated on 40 patients and was found to be 0.91 (116).

Validity. All AAOS instruments were developed by clinicians who continually confirmed the face and content validity for each questionnaire. Construct validity was obtained by correlating scores from the Hip and Knee Questionnaire with the AAOS Lower Limb Questionnaire (r = 0.95), the unweighted mean of three SF-36 physical subscales (r = 0.7), and a scale created from a physician assessment of function and pain (r = 0.73 and r = 0.69, respectively). In addition, a global score for the WOMAC was calculated and correlated with the AAOS Hip and Knee Questionnaire score (r = 0.89) (116).

Responsiveness. There is no direct measure of the Hip and Knee Questionnaire's responsiveness, but the Lower Limb Questionnaire (from which the Hip and Knee Questionnaire was adapted) has been assessed. An absolute change score, calculated as the difference between the baseline and follow-up scores, was found to be moderately correlated with a transition score (r = 0.53). This transition score was calculated from the combined responses of the patient and a physician on their perceptions of improvement over 1 year. Based on the strong correlation of Lower Limb Questionnaire scores to Hip and Knee Questionnaire scores, it is likely that the Hip and Knee Questionnaire possesses similar ability to detect change.

Minimally important differences. To our knowledge, no minimally important difference for the AAOS Hip and Knee Questionnaire has been calculated.

Generalizability. The AAOS Hip and Knee Questionnaire is applicable to adult populations older than age 18 with pain attributable to the hip or knee.

Use in clinical trials. To our knowledge, the AAOS Hip and Knee Questionnaire has not been used in any major clinical trials. It has been used to assess outcomes in patients with slipped capital femoral epiphysis (118), and the AAOS Lower Limb Questionnaire appears in numerous studies covering a broad range of musculoskeletal topics, including outcomes of limb lengthening and lower - extremity amputation (119,120).

Critical appraisal of overall value to the rheumatology community

Strengths. The AAOS Hip and Knee Questionnaire proved to be an effective instrument in assessing hip and knee conditions and the effect of appropriate treatments. It is a short survey that is easily administered and scored. A thorough psychometric evaluation was performed in 2004 to demonstrate its reliability and validity, and in conjunction with the SF-36, it can be an overall effective outcome measure.

Caveats and cautions. As patient-reported outcomes have become an integral part of the overall effort to deliver quality care in the United States, other surveys specific to hip and knee outcomes have been developed and endorsed by the Centers for Medicare and Medicaid Services (121). The AAOS Hip and Knee Questionnaire is not included in that group of endorsed outcome measures. This is likely the reason this survey is seen less commonly in the literature and why there have not been any follow-up studies to the original evaluations of the AAOS questionnaires published in 2004.

Clinical and research usability. The AAOS Hip and Knee Questionnaire is a useful tool given its size and ease of administration but is not practical given the rising popularity of other surveys, as mentioned above.

WESTERN ONTARIO AND MCMASTER UNIVERSITIES OSTEOARTHRITIS INDEX

Description

Purpose. The WOMAC was developed to measure the symptoms and physical disability of patients with hip or knee arthritis (122). It was created as part of a randomized controlled trial of two anti-inflammatory medications in the treatment of hip

and knee arthritis (123). As such, the WOMAC was intended to detect clinically important and relevant changes to treatment (124,125).

Content. The survey consists of three subscales that evaluate pain during certain positions or movements, stiffness at different times of day, and difficulty with performing certain activities.

Number of items. The WOMAC has 24 total items. There are 5 items in the pain section, 2 items in the stiffness section, and 17 items in the functional section.

Response options/scale. The WOMAC is available in a 5-point Likert scale, a 100-mm VAS, and an 11-box numerical rating scale (126). The Likert scale offers five response options ranging from "none" to "extreme," with corresponding score values from 0 to 4. The VAS and numerical rating scale offer responses selected on 100-mm or 11-box horizontal scales ranging from "none" on the left to "extreme" on the right (127).

Recall period for items. Forty-eight hours.

Cost to use. There is a fee to access the questionnaire and user guide. Fee information is available after submitting a request to use.

How to obtain. Requests to use the WOMAC for clinical or research purposes must be submitted via the contact section on the website (www.womac.com).

Practical application

Method of administration. The WOMAC is primarily designed as a self- or interview-administered questionnaire. The WOMAC has been validated for use in person, over the phone, or via computer or mobile phone (128–130). Patients have been shown to respond similarly to paper or online versions (131).

Scoring. For the Likert scale version, the pain, stiffness, and function sections possess potential summative score totals of 20, 8, and 68, respectively. A global score is calculated by combining the three subscores (124). For the VAS version, respondents mark a point along the horizontal 100-mm line. This is measured in millimeters and totaled out of 2400 (132). Normative values were published from an Australian population in 2010 (133). In the case that two or more items are missing from the pain subscale, that both items are missing from the stiffness subscale, or that four or more items are missing from the functional subscale, responses are declared invalid and not included for analysis.

Score interpretation. On both the Likert and VAS versions, lower scores equate to less pain and disability.

Respondent time to complete. Five to 10 minutes.

Administrative burden. Questionnaire responses require approximately 5 minutes to score, with minimal training required.

Translations/adaptations. There have been multiple versions of the WOMAC questionnaire, with the most recent being version 3.1 (updated in 2016). Short-form versions of the WOMAC have been developed, specific to total hip or total knee replacement, and were found to be equally responsive compared with the original WOMAC (134). The WOMAC has validated language translations in Arabic, Chinese, Dutch, Finnish, German, Hebrew, Italian, Japanese, Korean, Moroccan, Singaporean Spanish, Swedish, Thai, and Turkish (107,108,135–147).

Psychometric information

Floor and ceiling effects. Floor effects for WOMAC subscales or total scores are generally minimal or 0 (148). However, in one 2005 study, floor effects at 6 months and 2 years following THA were significant in the pain (25% and 39%) and stiffness (30% and 46%) subscales (149). In a study of patients after periacetabular osteotomy, floor effects in the pain and stiffness subscales were less than 1% and 3%, respectively. Floor effects were absent in the function subscale and aggregate score (7).

Ceiling effects are more commonly reported in all WOMAC subscales, particularly in pain and stiffness (7,95,148). A ceiling effect has been demonstrated in the pain and stiffness subscales when patients are evaluated at 10 weeks and 12 months after hip fracture (148). These effects appear to be significantly different between young (18%-36%) and old (38%-53%) age groups (148). In patients after periacetabular osteotomy, the ceiling effect is more substantial than the floor effect but still small (4%, 11%, 5%, and 2% for pain, stiffness, function, and total scores, respectively) (7). Findings in the literature are mixed; one study demonstrated low ceiling effects of less than 5% for all WOMAC subscales at 6 months and 2 years after THA, whereas another demonstrated significant ceiling effects (14%-38%) 1 year postoperatively (95,149).

Reliability. The WOMAC has demonstrated excellent internal and test-retest reliability. The Cronbach's α coefficient for the global score is reported as high as 0.97, with the lowest subscale being the stiffness subscale (α = 0.86) (7,127). Test-retest reliability is equally as strong, with Pearson's correlation coefficients reported as high as 0.96 (127). In addition to patients with osteoarthritis, WOMAC scores have demonstrated good internal reliability among patients after hip fracture, with Cronbach's coefficients ranging from 0.83 to 0.98 and from 0.79 to 0.97 for young and old age groups, respectively (148).

Validity. Since its creation in the late 1980s, the WOMAC has been extensively validated for patients with hip and knee arthritis (122). Additional validation has occurred for each of the aforementioned translations. More recently, the WOMAC has demonstrated good construct validity when compared with other PROMs. There is strong correlation of total WOMAC and SF-12 scores in patients after hip fracture, ranging from 0.71 to 0.83 and from 0.75 to 0.90 for young and old age groups, respectively (148). Similar results have been published, correlating WOMAC and EQ-5D subscales (148). In patients after periacetabular osteotomy, Spearman's correlation coefficients between WOMAC subscales and the SF-12 physical subscale are moderate to strong ($P > 0.5$), except for the stiffness subscale ($P = 0.38$) (7).

Responsiveness. For patients with hip osteoarthritis, the SRMs calculated for the WOMAC are high, exceeding 1.0 (150,151). In 162 patients who underwent THA for osteoarthritis, there was a mean WOMAC score change of 29, an effect size of 1.84, and an SRM of 1.6 (152). SRM values at 6 months and 2 years postoperatively have been shown to continue increasing to 1.86 and 1.98, respectively (149). In addition, a statistically significant improvement in scores has been shown up to 2 years following total hip replacement (149). In patients who sustain hip fractures, responsiveness is moderate (SRM 0.66) and small (SRM 0.24) among patients younger than and older than 80 years old, respectively (148). These differences are likely due to the fact that WOMAC scores are lowest just before a hip replacement and are not likely to be obtained prior to a patient sustaining a hip fracture (148).

Minimally important differences. The MCID for patients who underwent THA for hip arthritis is 10.2 points (152). The MCID for patients who undergo periacetabular osteotomy for hip pain secondary to hip dysplasia is approximately 11 points for the total WOMAC score (7). A prospective cohort study of more than 1300 patients identified a minimal clinically important improvement of 7.9 on the WOMAC function subscale in patients with hip arthritis who initiated nonoperative treatment (153). In addition, mean changes of 9 to 12 mm (100-mm normalized VAS) on WOMAC scales were perceptible by patients with hip and knee osteoarthritis (154).

Generalizability. The WOMAC questionnaire has been applied across wide groups of adult populations. To our knowledge, no study has validated the WOMAC in pediatric populations. One active study validating WOMAC scores in a pediatric population with Perthes disease can be found in the recruitment phase on ClinicalTrials.gov.

Use in clinical trials. As one of the oldest PROMs validated for assessing the hip, the WOMAC has been used extensively in studies affecting many types of hip pathology and responses to treatment, including osteoarthritis, hip dysplasia, and hip trauma.

Critical appraisal of overall value to the rheumatology community

Strengths. The WOMAC is a historically significant and widely used PROM that is found commonly in the literature. It has validated short-form versions and has been translated into multiple languages. It also serves as the foundation for other PROMs, such as the HOOS. The WOMAC subscales may also be valuable in stratifying and more thoroughly evaluating data.

Caveats and cautions. The WOMAC is proprietary and, as such, is less accessible and less attractive to potential clinicians or researchers. In addition, with 24 items, it is longer than newer, nonproprietary PROMs, such as the HOOS-JR. The full WOMAC is also not included in the list of PROMs recognized by the Centers for Medicare and Medicaid Services. However, the HOOS, a derivative of the WOMAC, is included on that list. As is common to all functional assessments, some activities included on the functional subscale may be impossible for some patients with severe disease to complete. This may result in missing data.

Clinical and research usability. If purchased, the WOMAC is relevant to both clinicians and researchers alike. It is reliable, valid, and responsive to treatment. Its ubiquitous presence in hip and knee outcome literature makes further use appropriate and guarantees the ability to compare new findings with old data.

CONCLUSIONS

We reviewed seven of the most commonly used instruments in assessment of hip outcomes and function. There has been an extensive body of work in terms of evaluation of psychometric properties and use in clinical trials since the last review of these instruments (1). This review should be used as a reference when comparing hip outcome measures and deciding which measures should be used for research or clinical purposes for researchers and clinicians.

AUTHOR CONTRIBUTIONS

All authors drafted the article, revised it critically for important intellectual content, and approved the final version to be published.

REFERENCES

1. Nilsdotter 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 (Hoboken) 2011;63 Suppl 11:S200–7.

2. 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.
3. McDonough CM, Stoiber E, Tomek IM, Ni P, Kim YJ, Tian F, et al. Sensitivity to change of a computer adaptive testing instrument for outcome measurement after hip and knee arthroplasty and periacetabular osteotomy. J Orthop Sports Phys Ther 2016;46:756–7.
4. Byrd JW, Jones KS. Prospective analysis of hip arthroscopy with 2-year follow-up. Arthroscopy 2000;16:578–87.
5. Wamper KE, Sierevelt IN, Poolman RW, Bhandari M, Haverkamp D. The Harris Hip Score: do ceiling effects limit its usefulness in orthopedics? [Systematic review]. Acta Orthop 2010;81:703–7.
6. Hung M, Hon SD, Cheng C, Franklin JD, Aoki SK, Anderson MB, et al. Psychometric evaluation of the lower extremity computerized adaptive test, the modified Harris Hip Score, and the Hip Outcome Score. Orthop J Sports Med 2014;2:2325967114562191.
7. Wasko MK, Yanik EL, Pascual-Garrido C, Clohisy JC. Psychometric properties of patient-reported outcome measures for periacetabular osteotomy. J Bone Joint Surg Am 2019;101:e21.
8. Söderman P, Malchau H. Is the Harris Hip Score system useful to study the outcome of total hip replacement? [Prospective cohort study]. Clin Orthop Relat Res 2001;189–97.
9. Kirmit L, Karatosun V, Unver B, Bakirhan S, Sen A, Gocen Z. The reliability of hip scoring systems for total hip arthroplasty candidates: assessment by physical therapists. Clin Rehabil 2005;19:659–61.
10. Kumar P, Sen R, Aggarwal S, Agarwal S, Rajnish RK. Reliability of modified Harris Hip Score as a tool for outcome evaluation of total hip replacements in Indian population. J Clin Orthop Trauma 2019;10:128–30.
11. Garellick G, Malchau H, Herberts P. Specific or general health outcome measures in the evaluation of total hip replacement: a comparison between the Harris Hip Score and the Nottingham Health Profile. J Bone Joint Surg Br 1998;80:600–6.
12. Shi HY, Mau LW, Chang JK, Wang JW, Chiu HC. Responsiveness of the Harris Hip Score and the SF-36: five years after total hip arthroplasty. Qual Life Res 2009;18:1053–60.
13. Frihagen F, Grotle M, Madsen JE, Wyller TB, Mowinckel P, Nordsletten L. Outcome after femoral neck fractures: a comparison of Harris Hip Score, Eq-5D and Barthel Index. Injury 2008;39:1147–56.
14. Ruecker AH, Rupprecht M, Gruber M, Gebauer M, Barvencik F, Briem D, et al. The treatment of intertrochanteric fractures: results using an intramedullary nail with integrated cephalocervical screws and linear compression. J Orthop Trauma 2009;23:22–30.
15. Dolatowski FC, Frihagen F, Bartels S, Opland V, Šaltytė Benth J, Talsnes O, et al. Screw fixation versus hemiarthroplasty for nondisplaced femoral neck fractures in elderly patients: a multicenter randomized controlled trial. J Bone Joint Surg Am 2019;101:136–44.
16. Barenius B, Inngul C, Alagic Z, Enocson A. A randomized controlled trial of cemented versus cementless arthroplasty in patients with a displaced femoral neck fracture: a four-year follow-up. Bone Joint J 2018;100-B:1087–93.
17. Doria C, Mosele GR, Caggiari G, Puddu L, Ciurlia E. Treatment of early hip osteoarthritis: ultrasound-guided platelet rich plasma versus hyaluronic acid injections in a randomized clinical trial. Joints 2017;5:152–5.
18. Winther SB, Foss OA, Husby OS, Wik TS, Klaksvik J, Husby VS. Muscular strength and function after total hip arthroplasty performed with three different surgical approaches: one-year follow-up study. Hip Int 2019;29:405–11.
19. Zomar BO, Bryant D, Hunter S, Howard JL, Vasarhelyi EM, Lanting BA. A randomised trial comparing spatio-temporal gait parameters after total hip arthroplasty between the direct anterior and direct lateral surgical approaches. Hip Int 2018;28:478–84.
20. Atrey A, Wolfstadt JI, Hussain N, Khoshbin A, Ward S, Shahid M, et al. The ideal total hip replacement bearing surface in the young patient: a prospective randomized trial comparing alumina ceramic-on-ceramic with ceramic-on-conventional polyethylene: 15-year follow-up. J Arthroplasty 2018;33:1752–6.
21. Fitzpatrick J, Bulsara MK, O'Donnell J, McCrory PR, Zheng MH. The effectiveness of platelet-rich plasma injections in gluteal tendinopathy: a randomized, double-blind controlled trial comparing a single platelet-rich plasma injection with a single corticosteroid injection. Am J Sports Med 2018;46:933–9.
22. Tol MC, van den Bekerom MP, Sierevelt IN, Hilverdink EF, Raaymakers EL, Goslings JC. Hemiarthroplasty or total hip arthroplasty for the treatment of a displaced intracapsular fracture in active elderly patients: 12-year follow-up of randomised trial. Bone Joint J 2017;99-B:250–4.
23. Makki D, Mohamed AM, Gadiyar R, Patterson M. Addition of an anti-rotation screw to the dynamic hip screw for femoral neck fractures. Orthopedics 2013;36:e865–8.
24. Madan SS, Cooper AP, Davies AG, Fernandes JA. The treatment of severe slipped capital femoral epiphysis via the Ganz surgical dislocation and anatomical reduction: a prospective study. Bone Joint J 2013;95-B:424–9.
25. Larson CM, Giveans MR, Stone RM. Arthroscopic debridement versus refixation of the acetabular labrum associated with femoroacetabular impingement: mean 3.5-year follow-up. Am J Sports Med 2012;40:1015–21.
26. Söderman P. On the validity of the results from the Swedish National Total Hip Arthroplasty Register. Acta Orthop Scand Suppl 2000;71:1–33.
27. Hoeksma HL, van den Ende CH, Ronday HK, Heering A, Breedveld FC. Comparison of the responsiveness of the Harris Hip Score with generic measures for hip function in osteoarthritis of the hip. Ann Rheum Dis 2003;62:935–8.
28. Klässbo 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.
29. Lyman S, Lee YY, Franklin PD, Li W, Mayman DJ, Padgett DE. Validation of the HOOS, JR: a short-form hip replacement survey. Clin Orthop Relat Res 2016;474:1472–82.
30. Davis AM, Perruccio AV, Canizares M, Tennant A, Hawker GA, Conaghan PG, et al. The development of a short measure of physical function for hip OA HOOS-Physical Function Shortform (HOOS-PS): an OARSI/OMERACT initiative. Osteoarthritis Cartilage 2008;16:551–9.
31. Nilsdotter AK, Lohmander LS, Klässbo M, Roos EM. Hip Disability And Osteoarthritis Outcome Score (HOOS): validity and responsiveness in total hip replacement. BMC Musculoskelet Disord 2003;4:10.
32. De Groot IB, Reijman M, Terwee CB, Bierma-Zeinstra SM, Favejee M, Roos EM, et al. Validation of the Dutch version of the Hip Disability And Osteoarthritis Outcome Score. Osteoarthritis Cartilage 2007;15:104–9.
33. Ornetti P, Parratte S, Gossec L, Tavernier C, Argenson JN, Roos EM, et al. Cross-cultural adaptation and validation of the French version of the Hip Disability And Osteoarthritis Outcome Score (HOOS) in hip osteoarthritis patients. Osteoarthritis Cartilage 2010;18:522–9.
34. 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.

35. Bini SA, Shah RF, Bendich I, Patterson JT, Hwang KM, Zaid MB. Machine learning algorithms can use wearable sensor data to accurately predict six-week patient-reported outcome scores following joint replacement in a prospective trial. J Arthroplasty 2019;34:2242–7.
36. Busch A, Jäger M, VITAS group, Wegner A, Haversath M. Vitamin E-blended versus conventional polyethylene liners in prostheses: prospective, randomized trial with 3-year follow-up. Orthopade 2019. E-pub ahead of print.
37. Otero JE, Martin JR, Rowe TM, Odum SM, Mason JB. Radiographic and clinical outcomes of modular tapered fluted stems for femoral revision for Paprosky III and IV femoral defects or Vancouver B2 and B3 femoral fractures. J Arthroplasty 2020;35:1069–73.
38. Lerch TD, Steppacher SD, Liechti EF, Tannast M, Siebenrock KA. One-third of hips after periacetabular osteotomy survive 30 years with good clinical results, no progression of arthritis, or conversion to THA. Clin Orthop Relat Res 2017;475:1154–68.
39. Li DJ, Clohisy JC, Schwabe MT, Yanik EL, Pascual-Garrido C. PROMIS versus legacy patient-reported outcome measures in patients undergoing surgical treatment for symptomatic acetabular dysplasia. Am J Sports Med 2020;48:385–94.
40. Flores SE, Sheridan JR, Borak KR, Zhang AL. When do patients improve after hip arthroscopy for femoroacetabular impingement? A prospective cohort analysis. Am J Sports Med 2018;46: 3111–8.
41. Padilla JA, Gabor JA, Schwarzkopf R, Davidovitch RI. A novel opioid-sparing pain management protocol following total hip arthroplasty: effects on opioid consumption, pain severity, and patient-reported outcomes. J Arthroplasty 2019;34:2669–75.
42. Dietz MJ, Klein AE, Lindsey BA, Duncan ST, Eicher JM, Gillig JD, et al. Posterior hip precautions do not impact early recovery in total hip arthroplasty: a multicenter, randomized, controlled study. J Arthroplasty 2019;34 Suppl:S221–7.
43. Banerjee S, Plummer O, Abboud JA, Deirmengian GK, Levicoff EA, Courtney PM. Accuracy and validity of computer adaptive testing for outcome assessment in patients undergoing total hip arthroplasty. J Arthroplasty 2020;35:756–61.
44. Westermann RW, Lynch TS, Jones MH, Spindler KP, Messner W, Strnad G, et al. Predictors of hip pain and function in femoroacetabular impingement: a prospective cohort analysis. Orthop J Sports Med 2017;5:2325967117726521.
45. Peters RM, van Beers LW, van Steenbergen LN, Wolkenfelt J, Ettema HB, Ten Have BL, et al. Similar superior patient-reported outcome measures for anterior and posterolateral approaches after total hip arthroplasty: postoperative patient-reported outcome measure improvement after 3 months in 12,774 primary total hip arthroplasties using the anterior, anterolateral, straight lateral, or posterolateral approach. J Arthroplasty 2018;33:1786–93.
46. Gagnier JJ, Huang H, Mullins M, Marinac-Dabić D, Ghambaryan A, Eloff B, et al. Measurement properties of patient-reported outcome measures used in patients undergoing total hip arthroplasty: a systematic review. JBJS Rev 2018;6:e2.
47. Veenhof C, Bijlsma JW, van den Ende CH, van Dijk GM, Pisters MF, Dekker J. Psychometric evaluation of osteoarthritis questionnaires: a systematic review of the literature. Arthritis Rheum 2006;55:480–92.
48. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. J Clin Epidemiol 2010;63:1179–94.
49. Ho B, Houck JR, Flemister AS, Ketz J, Oh I, DiGiovanni BF, et al. Preoperative PROMIS scores predict postoperative success in foot and ankle patients. Foot Ankle Int 2016;37:911–8.
50. Hung M, Stuart AR, Higgins TF, Saltzman CL, Kubiak EN. Computerized adaptive testing using the PROMIS physical function item bank reduces test burden with less ceiling effects compared with the short musculoskeletal function assessment in orthopaedic trauma patients. J Orthop Trauma 2014;28:439–43.
51. Hung M, Saltzman CL, Greene T, Voss MW, Bounsanga J, Gu Y, et al. Evaluating instrument responsiveness in joint function: the HOOS JR, the KOOS JR, and the PROMIS PF CAT. J Orthop Res 2018;36:1178–84.
52. Stiegel KR, Lash JG, Peace AJ, Coleman MM, Harrington MA, Cahill CW. Early experience and results using Patient-Reported Outcomes Measurement Information System scores in primary total hip and knee arthroplasty. J Arthroplasty 2019;34:2313–8.
53. Beleckas CM, Guattery J, Chamberlain AM, Khan T, Kelly MP, Calfee RP. Using Patient-Reported Outcomes Measurement Information System measures to understand the relationship between improvement in physical function and depressive symptoms. J Am Acad Orthop Surg 2018;26:e511–8.
54. Lawrie CM, Abu-Amer W, Barrack RL, Clohisy JC. Is the Patient-Reported Outcome Measurement Information System feasible in bundled payment for care improvement in total hip arthroplasty patients? [Retrospective cohort study]. J Arthroplasty 2020;35:1179–85.
55. Broderick JE, Schneider S, Junghaenel DU, Schwartz JE, Stone AA. Validity and reliability of Patient-Reported Outcomes Measurement Information System (PROMIS) instruments in osteoarthritis. Arthritis Care Res (Hoboken) 2013;65:1625–33.
56. Martusiewicz A, Delagrammaticas D, Harold RE, Bhatt S, Beal MD, Manning DW. Anterior versus posterior approach total hip arthroplasty: patient-reported and functional outcomes in the early postoperative period. Hip Int 2019. E-pub ahead of print.
57. Austin DC, Torchia MT, Moschetti WE, Jevsevar DS, Keeney BJ. Patient outcomes after total hip arthroplasty in extreme elderly patients older than 80 years. Hip Int 2020;30:407–16.
58. Sheean AJ, Schmitz MR, Ward CL, Barrow AE, Tennent DJ, Roach CJ, et al. Assessment of disability related to femoroacetabular impingement syndrome by use of the Patient-Reported Outcome Measure Information System (PROMIS) and objective measures of physical performance. Am J Sports Med 2017;45:2476–82.
59. Nwachukwu BU, Beck EC, Chapman R, Chahla J, Okoroha K, Nho SJ. Preoperative performance of the PROMIS in patients undergoing hip arthroscopic surgery for femoroacetabular impingement syndrome. Orthop J Sports Med 2019;7:2325967119860079.
60. Livermore AT, Anderson LA, Anderson MB, Erickson JA, Peters CL. Correction of mildly dysplastic hips with periacetabular osteotomy demonstrates promising outcomes, achievement of correction goals, and excellent five-year survivorship. Bone Joint J 2019;101-B Suppl B:16–22.
61. 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.
62. 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.
63. 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.
64. Baker PN, van der Meulen JH, Lewsey J, Gregg PJ, National Joint Registry for England and Wales. The role of pain and function in determining patient satisfaction after total knee replacement: data from the National Joint Registry for England and Wales. J Bone Joint Surg Br 2007;89:893–900.
65. Dunbar MJ, Robertsson O, Ryd L, Lidgren L. Appropriate questionnaires for knee arthroplasty: results of a survey of 3600 patients

from the Swedish Knee Arthroplasty Registry. J Bone Joint Surg Br 2001;83:339–44.
66. Pynsent PB, Adams DJ, Disney SP. The Oxford hip and knee outcome questionnaires for arthroplasty. J Bone Joint Surg Br 2005;87:241–8.
67. Rothwell AG, Hooper GJ, Hobbs A, Frampton CM. An analysis of the Oxford hip and knee scores and their relationship to early joint revision in the New Zealand Joint Registry. J Bone Joint Surg Br 2010;92:413–8.
68. Uesugi Y, Makimoto K, Fujita K, Nishii T, Sakai T, Sugano N. Validity and responsiveness of the Oxford Hip Score in a prospective study with Japanese total hip arthroplasty patients. J Orthop Sci 2009;14:35–9.
69. Wylde V, Blom AW, Whitehouse SL, Taylor AH, Pattison GT, Bannister GC. Patient-reported outcomes after total hip and knee arthroplasty: comparison of midterm results. J Arthroplasty 2009;24:210–6.
70. Dawson J, Fitzpatrick R, Frost S, Gundle R, McLardy-Smith P, Murray D. Evidence for the validity of a patient-based instrument for assessment of outcome after revision hip replacement. J Bone Joint Surg Br 2001;83:1125–9.
71. Field RE, Cronin MD, Singh PJ. The Oxford Hip Scores for primary and revision hip replacement. J Bone Joint Surg Br 2005;87:618–22.
72. Yeoman TF, Clement ND, Macdonald D, Moran M. Recall of preoperative Oxford hip and knee scores one year after arthroplasty is an alternative and reliable technique when used for a cohort of patients. Bone Joint Res 2018;7:351–6.
73. Eibich P, Dakin HA, Price AJ, Beard D, Arden NK, Gray AM. Associations between preoperative Oxford hip and knee scores and costs and quality of life of patients undergoing primary total joint replacement in the NHS England: an observational study. BMJ Open 2018;8:e019477.
74. Matharu GS, McBryde CW, Robb CA, Pynsent PB. An analysis of Oxford hip and knee scores following primary hip and knee replacement performed at a specialist centre. Bone Joint J 2014;96-B:928–35.
75. Scott CE, Turnbull GS, Powell-Bowns MF, MacDonald DJ, Breusch SJ. Activity levels and return to work after revision total hip and knee arthroplasty in patients under 65 years of age. Bone Joint J 2018;100-B:1043–53.
76. Abdel Messih M, Naylor JM, Descallar J, Manickam A, Mittal R, Harris IA. Mail versus telephone administration of the Oxford knee and hip scores. J Arthroplasty 2014;29:491–4.
77. Kalairajah Y, Azurza K, Hulme C, Molloy S, Drabu KJ. Health outcome measures in the evaluation of total hip arthroplasties: a comparison between the Harris Hip Score and the Oxford Hip Score. J Arthroplasty 2005;20:1037–41.
78. Hamilton DF, Loth FL, MacDonald DJ, Giesinger K, Patton JT, Simpson AH, et al. Treatment success following joint arthroplasty: defining thresholds for the Oxford hip and knee scores. J Arthroplasty 2018;33:2392–7.
79. Hamilton DF, Giesinger JM, Patton JT, MacDonald DJ, Simpson AH, Howie CR, et al. Making the Oxford hip and knee scores meaningful at the patient level through normative scoring and registry data. Bone Joint Res 2015;4:137–44.
80. Beard DJ, Harris K, Dawson J, Doll H, Murray DW, Carr AJ, et al. Meaningful changes for the Oxford hip and knee scores after joint replacement surgery. J Clin Epidemiol 2015;68:73–9.
81. Neufeld ME, Masri BA. Can the Oxford knee and hip score identify patients who do not require total knee or hip arthroplasty? [Retrospective cohort study]. Bone Joint J 2019;101-B:23–30.
82. McMurray R, Heaton J, Sloper P, Nettleton S. Measurement of patient perceptions of pain and disability in relation to total hip replacement: the place of the Oxford Hip Score in mixed methods. Qual Health Care 1999;8:228–33.
83. Wylde V, Learmonth ID, Cavendish VJ. The Oxford Hip Score: the patient's perspective. Health Qual Life Outcomes 2005;3:66.
84. Delaunay C, Epinette JA, Dawson J, Murray D, Jolles BM. Cross-cultural adaptations of the Oxford-12 HIP score to the French speaking population. Orthop Traumatol Surg Res 2009;95:89–99.
85. Nourbakhsh M, Zarezadeh A, Shemshaki H, Etemadifar MR, Moezi M, Mazoochian F. Translation and cultural adaptation of the Oxford Hip Score for Iranian population. Int J Prev Med 2013;4:141–5.
86. Lee J, Ozdoganlar OB, Rabin Y. An experimental investigation on thermal exposure during bone drilling. Med Eng Phys 2012;34:1510–20.
87. Fitzpatrick R, Norquist JM, Jenkinson C, Reeves BC, Morris RW, Murray DW, et al. A comparison of Rasch with Likert scoring to discriminate between patients' evaluations of total hip replacement surgery. Qual Life Res 2004;13:331–8.
88. Fitzgerald AM, Gaston P, Wilson Y, Quaba A, McQueen MM. Long-term sequelae of fasciotomy wounds. Br J Plast Surg 2000;53:690–3.
89. Péchon PH, Butler K, Murphy G, Singer GC. The perils of PROMs: question 5 of the Oxford Hip Score is ambiguous to 10% of English-speaking patients: a survey of 135 patients. Hip Int 2019;29:299–302.
90. Ostendorf M, van Stel HF, Buskens E, Schrijvers AJ, Marting LN, Verbout AJ, et al. Patient-reported outcome in total hip replacement: a comparison of five instruments of health status. J Bone Joint Surg Br 2004;86:801–8.
91. Impellizzeri FM, Mannion AF, Naal FD, Leunig M. Validity, reproducibility, and responsiveness of the Oxford Hip Score in patients undergoing surgery for femoroacetabular impingement. Arthroscopy 2015;31:42–50.
92. Gosens T, Hoefnagels NH, de Vet RC, Dhert WJ, van Langelaan EJ, Bulstra SK, et al. The "Oxford Heup Score": the translation and validation of a questionnaire into Dutch to evaluate the results of total hip arthroplasty. Acta Orthop 2005;76:204–11.
93. Fitzpatrick R, Morris R, Hajat S, Reeves B, Murray DW, Hannen D, et al. The value of short and simple measures to assess outcomes for patients of total hip replacement surgery. Qual Health Care 2000;9:146–50.
94. Kwong E, Neuburger J, Black N. Agreement between retrospectively and contemporaneously collected patient-reported outcome measures (PROMs) in hip and knee replacement patients. Qual Life Res 2018;27:1845–54.
95. Garbuz DS, Xu M, Sayre EC. Patients' outcome after total hip arthroplasty: a comparison between the Western Ontario and McMaster Universities index and the Oxford 12-item hip score. J Arthroplasty 2006;21:998–1004.
96. Weel H, Lindeboom R, Kuipers SE, Vervest TM. Comparison between the Harris- and Oxford Hip Score to evaluate outcomes one-year after total hip arthroplasty. Acta Orthop Belg 2017;83:98–109.
97. Dawson J, Fitzpatrick R, Murray D, Carr A. Comparison of measures to assess outcomes in total hip replacement surgery. Qual Health Care 1996;5:81–8.
98. 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.
99. Lequesne M. Indices of severity and disease activity for osteoarthritis. Semin Arthritis Rheum 1991;20 Suppl 2:48–54.

100. Lequesne MG. The algofunctional indices for hip and knee osteoarthritis. J Rheumatol 1997;24:779–81.
101. Stucki G, Sangha O, Stucki S, Michel BA, Tyndall A, Dick W, et al. Comparison of the WOMAC (Western Ontario and McMaster Universities) osteoarthritis index and a self-report format of the self-administered Lequesne-Algofunctional index in patients with knee and hip osteoarthritis. Osteoarthritis Cartilage 1998;6:79–86.
102. Dawson J, Linsell L, Doll H, Zondervan K, Rose P, Carr A, et al. Assessment of the Lequesne index of severity for osteoarthritis of the hip in an elderly population. Osteoarthritis Cartilage 2005;13:854–60.
103. Dougados M, Gueguen A, Nguyen M, Berdah L, Lequesne M, Mazieres B, et al. Requirement for total hip arthroplasty: an outcome measure of hip osteoarthritis? [Prospective cohort study]. J Rheumatol 1999;26:855–61.
104. Grimmig H, Melzer C, Ludwig FJ, Daalmann HH. Using the Lequesne questionnaires for patients with hip and knee disabilities in daily routine to reflect the quality of outcome in rehabilitation. Z Orthop Ihre Grenzgeb 2002;140:452–9. In German.
105. Faucher M, Poiraudeau S, Lefevre-Colau MM, Rannou F, Fermanian J, Revel M. Algo-functional assessment of knee osteoarthritis: comparison of the test-retest reliability and construct validity of the WOMAC and Lequesne indexes. Osteoarthritis Cartilage 2002;10:602–10.
106. Lequesne MG, Samson M. Indices of severity in osteoarthritis for weight bearing joints. J Rheumatol Suppl 1991;27:16–8.
107. Basaran S, Guzel R, Seydaoglu G, Guler-Uysal F. Validity, reliability, and comparison of the WOMAC osteoarthritis index and Lequesne Algofunctional Index in Turkish patients with hip or knee osteoarthritis. Clin Rheumatol 2010;29:749–56.
108. Bae SC, Lee HS, Yun HR, Kim TH, Yoo DH, Kim SY. Cross-cultural adaptation and validation of Korean Western Ontario and McMaster Universities (WOMAC) and Lequesne osteoarthritis indices for clinical research. Osteoarthritis Cartilage 2001;9:746–50.
109. Espinosa-Cuervo G, Guillermin F, Rat AC, Duarte-Salazar C, Alemán-Hernández SI, Vergara-Álvarez Y, et al. Transculturalization and validation of a Spanish translation of the specific lower limb osteoarthritis and quality of life questionnaire AMICAL: arthrose des membres inférieurs et qualité de vie AMIQUAL. Reumatol Clin 2014;10:241–7.
110. Konstantinidis GA, Aletras VH, Kanakari KA, Natsis K, Bellamy N, Niakas D. Comparative validation of the WOMAC osteoarthritis and Lequesne Algofunctional Indices in Greek patients with hip or knee osteoarthritis. Qual Life Res 2014;23:539–48.
111. Nadrian H, Moghimi N, Nadrian E, Moradzadeh R, Bahmanpour K, Iranpour A, et al. Validity and reliability of the Persian versions of WOMAC osteoarthritis index and Lequesne Algofunctional Index. Clin Rheumatol 2012;31:1097–102.
112. Ludwig FJ, Melzer C, Grimmig H, Daalmann HH. Cross cultural adaptation of the Lequesne Algofunctional Indices for German speaking patients with osteoarthritis of the hip and the knee. Rehabilitation (Stuttg) 2002;41:249–57. In German.
113. Stucki G, Meier D, Stucki S, Michel BA, Tyndall AG, Elke R, et al. Evaluation of a German questionnaire version of the Lequesne coxand gonarthrosis indices. Z Rheumatol 1996;55:50–7. In German.
114. Santos JP, da Silva RA, Fernandes MT, Poli-Frederico RC, Santos DC, Andraus RA, et al. Use of the Lower Extremity Functional Scale (LEFS-Brazil) questionnaire compared to Lequesne Algofunctional Index for definition of knee and hip osteoarthritis severity. Rev Bras Reumatol Engl Ed 2017;57:274–7.
115. Theiler R, Sangha O, Schaeren S, Michel BA, Tyndall A, Dick W, et al. Superior responsiveness of the pain and function sections of the Western Ontario and McMaster Universities osteoarthritis index (WOMAC) as compared to the Lequesne-Algofunctional Index in patients with osteoarthritis of the lower extremities. Osteoarthritis Cartilage 1999;7:515–9.
116. Johanson NA, Liang MH, Daltroy L, Rudicel S, Richmond J. American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments: reliability, validity, and sensitivity to change. J Bone Joint Surg Am 2004;86:902–9.
117. Hunsaker FG, Cioffi DA, Amadio PC, Wright JG, Caughlin B. The American Academy of Orthopaedic Surgeons outcomes instruments: normative values from the general population. J Bone Joint Surg Am 2002;84:208–15.
118. DeLullo JA, Thomas E, Cooney TE, McConnell SJ, Sanders JO. Femoral remodeling may influence patient outcomes in slipped capital femoral epiphysis. Clin Orthop Relat Res 2007;457:163–70.
119. Kim SJ, Balce GC, Agashe MV, Song SH, Song HR. Is bilateral lower limb lengthening appropriate for achondroplasia? Midterm analysis of the complications and quality of life. Clin Orthop Relat Res 2012;470:616–21.
120. Polfer EM, Hoyt BW, Bevevino AJ, Forsberg JA, Potter BK. Knee disarticulations versus transfemoral amputations: functional outcomes. J Orthop Trauma 2019;33:308–11.
121. Centers for Medicare and Medicaid Services. Measure methodology. URL: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Measure-Methodology.
122. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833–40.
123. Bellamy N, Buchanan WW, Grace E. Double-blind randomized controlled trial of isoxicam vs piroxicam in elderly patients with osteoarthritis of the hip and knee. Br J Clin Pharmacol 1986;22 Suppl 2:149S–55S.
124. McConnell S, Kolopack P, Davis AM. The Western Ontario and McMaster Universities osteoarthritis index (WOMAC): a review of its utility and measurement properties. Arthritis Rheum 2001;45:453–61.
125. Bellamy N. Outcome measurement in osteoarthritis clinical trials. J Rheumatol Suppl 1995;43:49–51.
126. Ornetti P, Dougados M, Paternotte S, Logeart I, Gossec L. Validation of a numerical rating scale to assess functional impairment in hip and knee osteoarthritis: comparison with the WOMAC function scale. Ann Rheum Dis 2011;70:740–6.
127. 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.
128. Bellamy N, Campbell J, Stevens J, Pilch L, Stewart C, Mahmood Z. Validation study of a computerized version of the Western Ontario and McMaster Universities VA3.0 osteoarthritis index. J Rheumatol 1997;24:2413–5.
129. Bellamy N, Wilson C, Hendrikz J, Whitehouse SL, Patel B, Dennison S, et al. Osteoarthritis index delivered by mobile phone (m-WOMAC) is valid, reliable, and responsive. J Clin Epidemiol 2011;64:182–90.
130. Theiler R, Spielberger J, Bischoff HA, Bellamy N, Huber J, Kroesen S. Clinical evaluation of the WOMAC 3.0 OA Index in numeric rating

scale format using a computerized touch screen version. Osteoarthritis Cartilage 2002;10:479–81.

131. Marsh JD, Bryant DM, Macdonald SJ, Naudie DDR. Patients respond similarly to paper and electronic versions of the WOMAC and SF-12 following total joint arthroplasty. J Arthroplasty 2014;29:670–3.
132. Bruce B, Fries J. Longitudinal comparison of the Health Assessment Questionnaire (HAQ) and the Western Ontario and McMaster Universities osteoarthritis index (WOMAC). Arthritis Rheum 2004;51:730–7.
133. Bellamy N, Wilson C, Hendrikz J. Population-based normative values for the Western Ontario and McMaster (WOMAC) osteoarthritis index and the Australian/Canadian (AUSCAN) hand osteoarthritis index functional subscales. Inflammopharmacology 2010;18:1–8.
134. Liebs TR, Herzberg W, Gluth J, Rüther W, Haasters J, Russlies M, et al. Using the patient's perspective to develop function short forms specific to total hip and knee replacement based on WOMAC function items. Bone Joint J 2013;95-B:239–43.
135. Guermazi M, Poiraudeau S, Yahia M, Mezganni M, Fermanian J, Habib Elleuch M, et al. Translation, adaptation and validation of the Western Ontario and McMaster Universities osteoarthritis index (WOMAC) for an Arab population: the Sfax modified WOMAC. Osteoarthritis Cartilage 2004;12:459–68.
136. Xie F, Li SC, Goeree R, Tarride JE, O'Reilly D, Lo NN, et al. Validation of Chinese Western Ontario and McMaster Universities osteoarthritis index (WOMAC) in patients scheduled for total knee replacement. Qual Life Res 2008;17:595–601.
137. Soininen JV, Paavolainen PO, Gronblad MA, Kaapa EH. Validation study of a Finnish version of the Western Ontario and McMasters University osteoarthritis index. Hip Int 2008;18:108–11.
138. Stucki G, Meier D, Stucki S, Michel BA, Tyndall AG, Dick W, et al. Evaluation of a German version of WOMAC (Western Ontario and McMaster Universities) arthrosis index. Z Rheumatol 1996;55:40–9. In German.
139. Wigler I, Neumann L, Yaron M. Validation study of a Hebrew version of WOMAC in patients with osteoarthritis of the knee. Clin Rheumatol 1999;18:402–5.
140. Salaffi F, Leardini G, Canesi B, Mannoni A, Fioravanti A, Caporali R, et al. Reliability and validity of the Western Ontario and McMaster Universities (WOMAC) osteoarthritis index in Italian patients with osteoarthritis of the knee. Osteoarthritis Cartilage 2003;11:551–60.
141. Faik A, Benbouazza K, Amine B, Maaroufi H, Bahiri R, Lazrak N, et al. Translation and validation of Moroccan Western Ontario and McMaster Universities (WOMAC) osteoarthritis index in knee osteoarthritis. Rheumatol Int 2008;28:677–83.
142. Hashimoto H, Hanyu T, Sledge CB, Lingard EA. Validation of a Japanese patient-derived outcome scale for assessing total knee arthroplasty: comparison with Western Ontario and McMaster Universities osteoarthritis index (WOMAC). J Orthop Sci 2003;8:288–93.
143. Thumboo J, Chew LH, Soh CH. Validation of the Western Ontario and Mcmaster University osteoarthritis index in Asians with osteoarthritis in Singapore. Osteoarthritis Cartilage 2001;9:440–6.
144. Escobar A, Quintana JM, Bilbao A, Azkárate J, Güenaga JI. Validation of the Spanish version of the WOMAC questionnaire for patients with hip or knee osteoarthritis: Western Ontario and McMaster Universities osteoarthritis index. Clin Rheumatol 2002;21:466–71.
145. Roos EM, Klässbo M, Lohmander LS. WOMAC osteoarthritis index: reliability, validity, and responsiveness in patients with arthroscopically assessed osteoarthritis: Western Ontario and MacMaster Universities. Scand J Rheumatol 1999;28:210–5.
146. Söderman P, Malchau H. Validity and reliability of Swedish WOMAC osteoarthritis index: a self-administered disease-specific questionnaire (WOMAC) versus generic instruments (SF-36 and NHP). Acta Orthop Scand 2000;71:39–46.
147. Kuptniratsaikul V, Rattanachaiyanont M. Validation of a modified Thai version of the Western Ontario and McMaster (WOMAC) osteoarthritis index for knee osteoarthritis. Clin Rheumatol 2007;26:1641–5.
148. Burgers PT, Poolman RW, van Bakel TM, Tuinebreijer WE, Zielinski SM, Bhandari M, et al. Reliability, validity, and responsiveness of the Western Ontario and McMaster Universities osteoarthritis index for elderly patients with a femoral neck fracture. J Bone Joint Surg Am 2015;97:751–7.
149. Quintana JM, Escobar A, Bilbao A, Arostegui I, Lafuente I, Vidaurreta I. Responsiveness and clinically important differences for the WOMAC and SF-36 after hip joint replacement. Osteoarthritis Cartilage 2005;13:1076–83.
150. Naal FD, Impellizzeri FM, von Eisenhart-Rothe R, Mannion AF, Leunig M. Reproducibility, validity, and responsiveness of the hip outcome score in patients with end-stage hip osteoarthritis. Arthritis Care Res (Hoboken) 2012;64:1770–5.
151. Bilbao A, Quintana JM, Escobar A, Las Hayas C, Orive M. Validation of a proposed WOMAC short form for patients with hip osteoarthritis. Health Qual Life Outcomes 2011;9:75.
152. SooHoo NF, Li Z, Chenok KE, Bozic KJ. Responsiveness of patient reported outcome measures in total joint arthroplasty patients. J Arthroplasty 2015;30:176–91.
153. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. Ann Rheum Dis 2005;64:29–33.
154. Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities osteoarthritis index questionnaire and global assessments in patients with osteoarthritis. J Rheumatol 2000;27:2635–41.

Table 1. Practical applications*

Measure	Number of Items	Content/ Domains	Method of Administration	Recall Period	Response Format	Range of Scores	Score Interpretation	Availability of Normative Data	Cross-Cultural Validation
HHS	10 items, 8 items in the modified HHS	Pain, function, absence of deformity, and ROM	Clinician-based; modified HHS is self-administered	Not described	Numerical scale within each domain	0-100	<70 indicates poor result, 70-80 indicates fair result, 80-90 indicates good result, 90-100 indicates excellent result	No	Used internationally but no validated versions in other languages
HOOS	40 items, HOOS-JR has 6 items, HOOS-PS has 5 items	Pain, other symptoms, function in ADLs, function in sport and recreation, and QOL	Self	Last week	5 Likert boxes	0-100	0 indicates extreme problems, 100 indicates no problems	No	25 total languages, including English, French, German, Dutch, Italian, and Spanish
PROMIS	4-12 items; Varies depending on exact measure chosen	Physical health, mental health, social health	Self	7 d; some questions present tense	0-10 for pain; 5 options for other domains	T-score metric with 50 as mean of population	50 is mean of population, 10 is 1 SD of population; 60 indicates 1 SD above mean	Yes	45 total languages, including English, French, German, Dutch, Italian, and Spanish
OHS	12 items	Pain, function	Self or over the phone	Last 4 wk	0-4 (worst to best)	0-48	0 indicates severe problems, 48 indicates no problems	No	English, Japanese, Dutch, German, Turkish, Spanish, Mandarin, and Italian; used, but not validated, in French, Iranian, and Korean
LISOH	11	Osteoarthritis-related symptoms, function	Self-administered, interviewer - based, or clinician-administered	Not specified	0-8 (best to worst)	0-28	0 indicates no handicap, >11-13 indicates severe handicap	No	English, French, German, Turkish, Korean, Spanish, Greek, Persian, and Portuguese
AAOS Hip and Knee Questionnaire	7	Pain, stiffness, swelling, function	Self	Last week	Likert scale for all domains with slight differences in options	0-100	0 indicates most disability, 100 indicates no disability	Yes	English and Spanish
WOMAC	24	Pain, stiffness, difficulty with certain activities	Self-administered or interviewer-administered	Last 48 h	Likert scale, VAS, and 11-box numerical rating scales	Pain: 0-20; stiffness: 0-8; function: 0-68; VAS totaled to 2400	Lower scores indicate less pain and disability	Yes	English, Arabic, Chinese, Dutch, Finnish, German, Hebrew, Italian, Japanese, Korean, Moroccan, Singaporean Spanish, Swedish, Thai, and Turkish

* AAOS = American Academy of Orthopaedic Surgeons; ADL = activity of daily living; HHS = Harris Hip Score; HOOS = Hip Disability and Osteoarthritis Outcome Score; HOOS-JR = Hip Disability and Osteoarthritis Outcome Score–Joint Replacement Short Form; HOOS-PS = Hip Disability and Osteoarthritis Outcome Score–Physical Function Short Form; LISOH = Lequesne Index of Severity for Osteoarthritis of the Hip; OHS = Oxford Hip Score; PROMIS = Patient-Reported Outcomes Measurement Information System; QOL = quality of life; ROM = range of motion; VAS = visual analog scale; WOMAC = Western Ontario McMaster Universities Osteoarthritis Index.

Table 2. Psychometrics*

Measure	Floor, Ceiling Effects	Reliability	Validity	Responsiveness	Minimally Important Differences	Generalizability	Used in RCTs
HHS	Ceiling effects: approximately 20%; floor effects: not reported	Internal consistency: Cronbach's α score = 0.94; test-retest reliability: r = 0.93-0.98	Content and construct validity correlate with similar domains of other scores	Responsive to pain and function at 6-mo follow-up, weakly responsive at 2-y follow-up; modified HHS adequately responsive after PAO	SRM = 0.75	Widely generalizable to adult patients with hip disabilities or undergoing hip surgeries	Yes
HOOS	Floor effects: 4.1%-17.8%; ceiling effects: up to 19%	Internal consistency: Cronbach's α = 0.82 -0.98; test-retest reliability = 0.75-0.97	Content validity performed by asking patients to rate item importance; construct validity confirmed by correlation with other scores	Good responsiveness in short- and long-term follow-up up to 2 y postoperatively	9.0-9.6 in ADL domain, 16.2-19.0 in QOL domain	Widely generalizable to adult patients with hip disabilities or undergoing hip procedures; HOOS-JR developed for patients undergoing THA	Yes
PROMIS	Not observed in patients undergoing PAO; floor effects in patients undergoing THA: 20%-45% in depression domain, 21%-26% in pain domain; no ceiling effects in patients undergoing THA	Internal consistency: Cronbach's α = 1.00	Not evaluated	High responsiveness postoperatively in patients undergoing THA	5 (half of normalized SD of reported scores)	Generalizable to adult patients with hip ailments or undergoing hip procedures; separate PROMIS instruments available for pediatric patients	Overall, less frequently used in RCTs but some recent use
OHS	Ceiling effects: 13.5%; low floor effects	Internal consistency: Cronbach's α = 0.84-0.93	High correlation with content validity when compared with other PROMs	High responsiveness compared with generic measures	Estimated to be 3-5 points (0- to 48-point scale)	Widely generalizable to adult patients with hip disabilities or undergoing hip surgeries	Yes
LISOH	Not reported	Internal consistency: Cronbach's α = 0.83-0.94 for composite score, pain/discomfort lower at 0.63; test-retest reliability = 0.51-0.96	Construct and convergent validity inferior to other PROMs	Not reported	Unknown	Not recommended as sole outcome measure clinically given poor validity and reliability; may have some utility in larger patient populations	Yes
AAOS Hip and Knee Questionnaire	Not measured	Internal consistency: Cronbach's α = 0.8; test-retest reliability = 0.91	Construct validity with moderate to high correlation to other PROMs	Not directly measured	Unknown	Widely generalizable to adult patients with hip disabilities or undergoing hip surgeries	Little use in major clinical trials
WOMAC	Floor effects generally thought of as minimal but one report of significant floor effects in pain and stiffness after THA (25%-46%); ceiling effects more common up to 53%	Internal consistency: Cronbach's α = 0.86-0.97; test-retest reliability = 0.96	Good construct validity compared with other PROMs	High SRM (>1.0)	MCID after THA 10.2, after PAO 11, 7.9 in nonoperative management of hip OA (scale of 0-68)	Widely generalizable to adult patients with hip disabilities or undergoing hip surgeries; currently being assessed for validity in pediatric population	Yes

* AAOS = American Academy of Orthopaedic Surgeons; ADL = activity of daily living; HHS = Harris Hip Score; HOOS = Hip Disability and Osteoarthritis Outcome Score; HOOS-JR = Hip Disability and Osteoarthritis Outcome Score–Joint Replacement Short Form; HOOS-PS = Hip Disability and Osteoarthritis Outcome Score–Physical Function Short Form; LISOH = Lequesne Index of Severity for Osteoarthritis of the Hip; MCID = minimal clinically important difference; OA = osteoarthritis; OHS = Oxford Hip Score; PAO = periacetabular osteotomy; PROM = patient-reported outcome measure; PROMIS = Patient-Reported Outcomes Measurement Information System; QOL = quality of life; RCT = randomized controlled trial; SRM = standardized response mean; THA = total hip arthroplasty; VAS = visual analog scale; WOMAC = Western Ontario McMaster Universities Osteoarthritis Index.