Arthritis Care & Research
Vol. 72, No. S10, October 2020, pp 219–249
DOI 10.1002/acr.24235

American College of Rheumatology
Empowering Rheumatology Professionals

Measures of Adult Knee Function

Michael McHugh, Erin Droy, Stefano Muscatelli, and Joel J. Gagnier

INTRODUCTION

Patient-reported outcome measures (PROMs) are important to fulfill both clinical and research purposes with regard to assessing knee function in patients with a variety of knee conditions associated with injury, osteoarthritis, or rheumatological disorders. For inclusion in this review, measures of knee function were required to be pertinent to rheumatology, orthopedics, and sports medicine specialties. We identified measures published with scientific analysis and included dimensions that were most important to patients, including pain, quality of life (QOL), and activity level. A 2011 review of nine tools was published and focused on many of these issues as they related to rheumatology and orthopedic surgery.

Based on the aforementioned criterion and the goals of this review, we used the same nine measures developed specifically for patient-reported knee function and perceptions: the International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, the Knee Injury and Osteoarthritis Outcome Score (KOOS), the KOOS Physical Function Short Form (KOOS-PS), the Knee Outcome Survey Activities of Daily Living Scale (KOS-ADLS), the Lysholm Knee Scoring Scale (LKS), the Oxford Knee Score (OKS), the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), the Activity Rating Scale (ARS), and the Tegner Activity Score (TAS). This updated review was conducted by doing a thorough search of new publications for each measure between January 1, 2010, and March 1, 2020. We also included the Patient-Reported Outcomes Measurement Information System Physical Function (PROMIS-PF) measure based on its rising popularity and the amount of research dedicated for its use in a variety of knee conditions.

A basic summary of the properties of the different measures is displayed in Table 1. Psychometric data pertaining to the floor and ceiling effects, validity, reliability, responsiveness, and minimum clinically important difference (MCID) of each patient-reported outcome are displayed in Table 2. Floor and ceiling effects were considered to be absent if no participants scored the bottom or top score, respectively, and to be acceptable if less than 15% of the cohort scored the bottom or top score, respectively. Validity was measured by assessing content, face, and construct validity. Content validity was present if patients were involved in development. Face validity was present if expert reviewers made a similar assessment and considered the measured items adequate. Construct validity was considered adequate if expected correlations were found with existing measures that assess similar (convergent construct validity) and dissimilar (divergent construct validity) constructs. Internal consistency was considered adequate if Cronbach's α was at least 0.7 (1), and test-retest (intrarater) reliability was adequate if the intraclass correlation coefficient was at least 0.8 for groups and 0.9 for individuals. Responsiveness was determined with a measure of ability to detect change over a period of time or intervention. MCID is the amount of change in a patient-reported outcome that represents a meaningful change to the patient.

Since 2011, there have been numerous studies evaluating the psychometric properties of the above measures. There has also been ample research assessing the utility and psychometric properties of the PROMIS-PF function. Extensive work has been performed to add available translations and culturally adapted versions of the above measures. Our review summarizes the available information about how these measures perform for different patient populations in different settings.

INTERNATIONAL KNEE DOCUMENTATION COMMITTEE SUBJECTIVE KNEE EVALUATION FORM

Description

Purpose. The IKDC detects improvement or deterioration in symptoms, function, and sports activities due to knee impairment caused by a variety of knee conditions, including ligament injuries, meniscal injuries, articular cartilage lesions, and patellofemoral pain (2).

Version. The IKDC was formed in 1987 to develop a standardized international documentation system for knee conditions. The IKDC Standard Knee Evaluation Form, which was designed

Michael McHugh, MD, Erin Droy, MD, Stefano Muscatelli, 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, Domino's Farms, 24 Frank Lloyd Wright Drive, Ann Arbor, MI 48105. E-mail: jgagnier@med.umich.edu.

Submitted for publication March 16, 2020; accepted in revised form April 21, 2020.

for knee ligament injuries, was subsequently published in 1993 (3) and revised in 1994 (4). The IKDC Subjective Knee Evaluation Form was developed as a revision of the Standard Knee Evaluation Form in 1997. It has undergone subsequent minor revisions since its publication in 2001.

Content. The IKDC Subjective Knee Evaluation Form contains the following three domains: 1) symptoms, including pain, stiffness, swelling, locking/catching, and giving way; 2) sports and daily activities; and 3) current knee function and knee function prior to knee injury (not included in the total score) (2).

Number of items. The IKDC Subjective Knee Evaluation Form consists of 18 items (7 items for symptoms, 1 item for sport participation, 9 items for daily activities, and 1 item for current knee function).

Response options/scale. Response options vary for each item. Item 6 dichotomizes response into yes/no; items 1, 4, 5, 7, 8, and 9 use five-point Likert scales; and items 2, 3, and 10 use 11-point numerical rating scales.

Recall period for items. The recall period is not specified for items 1, 3, 5, 7, 8, and 9; it is 4 weeks for items 2, 4, and 6. Item 10a refers to function prior to knee injury, and item 10b refers to current function.

Cost to use. Free to use. Cost of administration and information storage was not assessed and varies for each practice.

How to obtain. See https://www.sportsmed.org/aossmimis/Staging/Research/IKDC_Forms.aspx.

Practical application

Method of administration. The IKDC Subjective Knee Evaluation Form is a patient-completed questionnaire. The form has not been validated for administration by interview, either in person or via telephone.

Scoring. The response to each item is scored using an ordinal method (ie, 0 for responses that represent the highest level of symptoms or lowest level of function). The most recent version has assigned scores for each possible response printed on the questionnaire. Scores for each item, excluding item 10a, are summed to give a total score. The total score is calculated as the sum of items divided by the maximum possible score multiplied by 100, to give a total score out of 100. An online scoring sheet is available (www.sportsmed.org/tabs/research/ikdc.aspx) that provides a patient's raw score and percentile score (relative to age- and sex-based norms). The item regarding knee function prior to knee injury is not included in the total score. The revised scoring method states that in cases in which patients have up to two missing values (ie, responses have been provided for at least 16 items), the total score is calculated as the sum of completed items divided by the maximum possible sum of completed items multiplied by 100.

Score interpretation. Possible scores range from 0 to 100, in which 100 = no limitation with daily or sporting activities and the absence of symptoms. Normative data are available from the general US population stratified for age, sex, and current/prior knee problems (5).

Respondent burden. The IKDC Subjective Knee Evaluation Form takes 10 minutes to complete (6). It uses simple language that is suitable for patients.

Administrative burden. The IKDC Subjective Knee Evaluation Form takes approximately 5 minutes to score. Training is not necessary. Manual scoring can be performed easily using the scoring instructions supplied with the questionnaire.

Translations/adaptations. The form is available in English, Arabic, Brazilian Portuguese, traditional Chinese (Taiwan and Hong Kong), simplified Chinese (China and Singapore), Czech, Dutch, French, German, Greek, Italian, Japanese, Korean, Norwegian, Polish, Spanish, Swedish, Thai, and Turkish. Cross-cultural adaptations have been conducted for the Arabic (7), Brazilian (8), Chinese (9–11), Danish (12), Dutch (13), German (14), Greek (15), Italian (6), Korean (16), Romanian (17), Thai (18), and Turkish (19) translations.

Psychometric information

Method of development. The initial set of items was developed by the IKDC by considering questions from the Standard Knee Evaluation Form, the Musculoskeletal Outcomes Data Evaluation and Management Systems Lower Limb Instrument, and the Activities of Daily Living (ADLs) and Sports Activity Scales of the Knee Outcome Survey. Pilot testing of the initial version (n = 144) resulted in revision or deletion of existing items and the addition of new items. Testing of the second version (n = 222) resulted in further revisions and deletions (based on missing data), producing a final version. Item-response theory was used to create the scoring system. Patients were not involved in development; rather, the items were selected by the IKDC, a committee of international orthopedic surgeons (2). Following development, validation, and implementation of the IKDC Subjective Knee Evaluation Form, a pediatric form was developed (the Pedi-IKDC), which has been tested for psychometric properties and normative data as well as electronic use (20–22).

Floor and ceiling effects. Studies consistently report no floor or ceiling effects (ie, no participants scored the lowest or highest score) (2,6,8,13,15,23,24).

Reliability. Internal consistency is adequate for patients with knee injuries and mixed knee pathologies (Table 1). Test-retest reliability is adequate for groups of patients with knee injuries and mixed pathologies and individuals with knee injuries. It has also been shown to be adequate in pediatric populations (20). The test-retest reliability is slightly below adequate for individuals who fall into a broader category of knee pathologies. However, studies have shown superior reliability over other measurement forms. The Chinese IKDC has better reliability than the Chinese KOOS (25). The Dutch IKDC had better reliability than the WOMAC and KOOS for meniscal injury (26).

Validity. *Face and content validity.* The domains covered by the IKDC Subjective Knee Evaluation Form appear to represent elements that are likely to be important to patients. However, the lack of patient contribution to the selection and revision of items in the IKDC Subjective Knee Evaluation Form means that content validity cannot necessarily be assumed.

Construct validity. There are consistent reports of high convergent and divergent construct validity, with the IKDC Subjective Knee Evaluation Form more strongly correlated with the Short Form 36 (SF-36) physical subscales and component summary than with the mental subscales and component summary (2,8,11,13,23,24,27). Construct validity is acceptable in the pediatric form (20) and improved over the KOOS-Child form (28). Studies have shown the IKDC Subjective Knee Evaluation Form score to be highly correlated with the Cincinnati Knee Rating System, pain visual analog scale (VAS), Oxford 12 Questionnaire, WOMAC, Lysholm score, and SF-36 physical component, physical function, and bodily pain subscales (8,13,29).

Responsiveness. The IKDC Subjective Knee Evaluation Form has been shown to be adequately responsive (24). In a study comparing responsiveness of the IKDC Subjective Knee Evaluation Form with that of the KOOS for anterior cruciate ligament (ACL) injuries, the IKDC Subjective Knee Evaluation Form was found to be adequately responsive, but the KOOS was not (30). The same finding was found in a Chinese study comparing the IKDC Subjective Knee Evaluation Form with the KOOS (25). Further testing has specifically shown its adequate responsiveness for meniscal injury (31). The Pedi-IKDC has also been shown to have acceptable responsiveness (20). When directly compared with the KOOS-Child, it has superior responsiveness (28).

Minimally important differences. The minimal detectable change has been reported to be between 8.8 and 15.6, and the SEM has been reported to be between 3.2 and 5.6. Few studies have shown the minimal important changes (MICs). One study shows the MIC to be 10.9 for meniscal injuries (31). Another study showed it to be 9.8 in the Chinese population (25), and another showed 12.0 for pediatric populations (12). The MCID has been reported to be 6.3 at 6 months and 16.7 at 12 months following cartilage repair (32) and 11.5 to 20.5 (range 6-28 months) in those who have undergone various surgical procedures for mixed (various) knee pathologies (33). The patient-acceptable symptom state (PASS) has not been determined.

Critical appraisal of overall value to the rheumatology community

Strengths. At face value, the domains covered by the IKDC Subjective Knee Evaluation Form appear to represent elements that are likely to be important to patients. It shows adequate internal consistency and has no floor or ceiling effects across mixed groups of patients with knee conditions. The IKDC Subjective Knee Evaluation Form has been shown to be responsive to change following surgical interventions, highlighting its usefulness in this patient population. It has particularly been shown to be a stronger measure for ACL injuries and meniscal injuries. It has also been shown to be a strong measure in the pediatric population.

Caveats and cautions. Despite demonstrating face validity, the lack of patient contribution to item selection indicates that content validity cannot necessarily be assumed and has not been thoroughly investigated. The relatively long recall period associated with three of the items may be a problem for some patients. The use of one aggregate score to represent symptoms, activities, and function may mask deficits in one domain. Psychometric testing is lacking for patients with knee osteoarthritis as an isolated group, as is responsiveness following nonsurgical management, highlighting areas for future studies.

Clinical usability. The IKDC Subjective Knee Evaluation Form involves minimal administrative and respondent burden and can be easily scored in the clinic using the online scoring sheet. However, clinicians using the online scoring system need to keep in mind that the normative data provided are from a particular population and may not be representative of their individual patient's population.

Research usability. Psychometric evaluation supports the use of the IKDC Subjective Knee Evaluation Form in research for a variety of knee conditions. Because some versions of the measure published online contain subtle differences in the wording of instructions and items, researchers should ensure that they utilize the version published as a component of the 2000 IKDC Knee Forms to ensure that findings of psychometric properties still apply and that comparisons can be made with previous studies. Administrative and respondent burden would not limit research use, although researchers should be diligent in checking for missing data.

KNEE INJURY AND OSTEOARTHRITIS OUTCOME SCORE

Description

Purpose. To measure the opinions of young, middle-aged, and elderly patients with posttraumatic osteoarthritis (OA) and other injuries leading to OA, regarding their knees and associated problems over short- and long-term follow-up (34). Examples of conditions include knee ligament injury (ACL, posterior cruciate ligament [PCL], or medial collateral ligament [MCL]), meniscal tears, knee cartilage lesions, knee OA, and osteochondritis dissecans. Interventions include ligament reconstruction (ACL, PCL, or MCL), meniscectomy, microfracture, osteochondral autografts, tibial osteotomy, total knee replacement (TKR), exercise (land-based or aquatic), intra-articular sodium hyaluronate injection, pharmacologic therapy, and glucosamine supplementation.

Content. The KOOS consists of the following five domains: 1) pain frequency and severity during functional activities; 2) symptoms such as the severity of knee stiffness and the presence of swelling, grinding or clicking, catching, and range of motion restriction; 3) difficulty experienced during ADLs; 4) difficulty experienced with sport and recreational activities; and 5) knee-related QOL (34). The original KOOS remains unchanged, although there have been other subscales developed, including the KOOS-12 Short Form, the KOOS-Joint Replacement Form, and the KOOS-Child Form

Number of items. The KOOS contains 42 items across five subscales.

Response options/scale. All items are rated on a five-point Likert scale (0-4) specific to each item.

Recall period for items. The recall period is the previous week for the pain, symptoms, ADL, and sport/recreation subscales. It is not defined for the QOL subscale.

Cost to use. Free of charge. The costs of distribution, collection, and data storage are not assessed.

How to obtain. The KOOS is available with associated documentation at www.koos.nu.

Practical application

Method of administration. The KOOS is a patient-completed, in-person questionnaire. It can be administered on paper, tablet, or computer.

Scoring. Scoring sheets (manual and computer spreadsheets) are provided on the website. Each item is scored from 0 to 4. The five dimensions are scored separately as the sum of all corresponding items. Scores are then transformed to a 0 to 100 scale (percentage of total possible score achieved) (34). If a mark is placed outside a box, the closest box is chosen. If two boxes are marked, the box that indicates more severe problems is chosen. One or two missing values within a subscale are substituted with the average value for that subscale. If two or more items are missing, the response is considered invalid, and a subscale score is not calculated.

Score interpretation. A score of 0 is equivalent to the most severe knee problems, and a score of 100 is representative of no knee problems. Population-based normative data stratified by age and sex are available (35).

Respondent time to complete. The KOOS takes 10 minutes to complete (34). It uses simple language and similar one-word responses for each item.

Administrative burden. The KOOS takes approximately 5 minutes to score if using the scoring spreadsheet. It can be automatically calculated with any type of data management software.

Translations/adaptations. The KOOS is available in Egyptian Arabic, Saudi Arabian Arabic, Austrian German, Bengali (India), Czech, Hong Kong Chinese, Singapore Chinese, Croatian, Danish, Dutch, Estonian, English, Finnish, Filipino (Philippines), French, German, Greek, Hindi (India), Icelandic, Italian, Japanese, Kannada (India), Korean, Latvian, Lithuanian, Malayalam (India), Malay, Marathi (India), Norwegian, Persian, Portuguese, Brazilian Portuguese, Polish, Romanian, Russian, Singapore, Slovakian, Slovenian, Spanish, US Spanish, Peruvian Spanish, Swedish, Tamil (India), Telugu (India), Thai, Turkish, Ukrainian, Urdu (India), Vietnamese, Welsh, and Zulu. Validation of the cross-cultural adaptations have been conducted and found adequate in the following languages (36,37): mainland Chinese (38, 39), Singapore Chinese (40), Greek (41), Icelandic (42), Spanish (43,44), Dutch (45), French (46), Saudi Arabian Arabic (47), Hong Kong Chinese (48), Japanese (49), Persian (50), Portuguese (51), Russian, Singapore English (40), Thai (52), and Turkish (53).

Psychometric information

Floor and ceiling effects. Studies consistently report no floor or ceiling effects or acceptable floor or ceiling effects in knee injury cohorts (36,50,54) and in patients with mild or moderate knee OA (37,40,46,51). In those with severe OA awaiting TKR (37,40,45,46,51), there are consistent reports of floor effects for the sport/recreation subscale (16%-73.3% scored the lowest score), and ceiling effects have been reported for the pain (15%-22%), sport/recreation (16%), and QOL (17%) subscales for up to 12 months following TKR (37). Comparatively, it has been shown to have lower ceiling effects in all categories except for pain

against the Knee Society Function score (55). Studies have shown that the original KOOS was not well understood by children, and subsequently the KOOS-Child was formed (56). The KOOS-Child was found to have no floor or ceiling effects (57).

Reliability. For patients with knee injuries, the pain, ADL, and sport/recreation subscales have adequate internal consistency in all reports, whereas the symptom and QOL subscales have had reports of lower or adequate internal consistency (Table 1). In patients with knee OA, the ADL, sport/recreation, and QOL subscales have adequate internal consistency, whereas the pain and symptoms subscales have reports of lower or adequate internal consistency. Test-retest reliability is adequate for group evaluation in all reports on the pain, symptoms, and QOL subscales for patients with knee injuries, whereas there are reports of lower and adequate reliability, respectively, for the ADL and sport/recreation subscales. Recent meta-analysis has shown adequate test-retest reliability for age- and condition-relevant subscales (58). Across the five subscales, the minimal detectable change ranges from 6 to 12 for knee injuries and from 13.4 to 21.1 for knee OA. For the five KOOS subscales, the pooled smallest detectable change (SDC) for individuals ranged from 15.7 (ADL) to 25.1 (sport/recreation). The SDC was greater for older adults and those with knee OA than for younger and ACL cohorts (58). The SEM is reported to be lower for knee injuries than for OA.

Validity. *Face and content validity.* In addition to exhibiting face validity, the KOOS shows content validity facilitated by direct involvement of patients with knee conditions in the development of the KOOS (34,37).

Construct validity. Multiple studies report that the KOOS demonstrates convergent and divergent construct validity, with the KOOS more strongly correlated with subscales of the SF-36 that measure similar constructs (eg, ADL with physical function, sport/recreation with physical function, and pain with bodily pain) and less strongly with SF-36 subscales that measure mental health (34,36,37,40,45,50,51,54,58,59). Rasch analysis conducted using patient data 20 weeks post–ACL reconstruction showed that only the sport/recreation and QOL subscales exhibited unidimensionality and that the three subscales that were based on the WOMAC did not (60). A more recent study reported that the KOOS subscales had acceptable dimensionality (59). Further meta-analysis more recently found the hypothesis of superior convergent and divergent construct validity were supported when all data were pooled and when data were split by age group and knee condition for the pain, symptoms, ADL, sports/recreation, and QOL subscales (58). They found that further testing was necessary for the short form as well as for structural validity in all categories (58).

Responsiveness. The KOOS appears to be responsive to change in patients with a variety of conditions that have been treated with nonsurgical and surgical interventions (Table 2). In patients who have undergone partial meniscectomy 3 months previously, large effect sizes are seen on all but the ADL subscale. Large effect sizes are seen in all subscales 6 months after ACL reconstruction. Three years following autologous chondrocyte implantation or microfracture, large effect sizes are seen for the pain, sport/recreation, and QOL subscales, and moderate effects are seen for the symptoms and ADL subscales. In those with knee OA who have undergone physical therapy treatment, large effect sizes are seen at 4 weeks on the pain, symptoms, and ADL subscales, whereas the sport/recreation and QOL subscales show moderate effects. Larger effect sizes are found following TKR than nonoperative treatment (58). Large effect sizes are consistently reported on all subscales 3 to 12 months after TKR, but the effect sizes do not increase over these periods (58). Large effect sizes have been shown to be a strength of the KOOS as opposed to other parameters (61).

Minimally important differences. The MCID of the KOOS Short Form and KOOS QOL has been reported in one study (62). MCID and moderate improvement estimates for the KOOS QOL were 8.0 and 15.6, respectively (62).

Critical appraisal of overall value to the rheumatology community

Strengths. The KOOS has undergone a substantial amount of psychometric testing. Over the last decade, the creation of subscales paired with psychometric testing has expanded as has cultural adaptation testing. Establishment of the KOOS as a reliable and valid measure across multiple languages highlights its usefulness as a patient-reported measure of knee function for people with knee OA and various combinations of sports- and trauma-related injuries. This has been expanded to include a child form of the test. The use of individual scores for each subscale, rather than an aggregate score, enhances clinical interpretation and, in research, acknowledges the impact of different interventions on different dimensions (eg, exercise therapy is likely to have more impact on ADLs and sports/recreation, whereas pharmacology may have more impact on pain and symptoms) and ensures content validity in groups of different ages and functional activity levels (eg, the sport/recreation subscale is more important in patients with a high physical activity level, whereas the ADL subscale is more important in subjects with a lower physical activity level).

Caveats and cautions. The KOOS has not been validated for interview administration, meaning that it may not be appropriate for patients who are unable to read or write or in cases when telephone follow-up is necessary. When administering the KOOS in older or less physically active individuals, higher-level components of the ADL and sport/recreation subscales may not be applicable and could result in missing data. It may be appropriate to leave out

the sport/recreation subscale in those with more advanced disease or disability; however, doing so omits the ability to measure improvements seen in these more demanding functions following treatment (37). The MCID has been minimally examined.

Clinical usability. The KOOS is freely available online. Administration and scoring burden are minimal when online score sheets are utilized. Clinicians should bear in mind that the sport/recreation subscale may not be applicable for less physically active patients and may not have adequate test-retest reliability in individuals with knee injuries.

Research usability. The KOOS fulfills desired criteria for research outcomes, demonstrating adequate reliability for use in groups and validity when used in those with knee injuries and knee OA. The inclusion of the three WOMAC subscales facilitates the comparison of findings with studies that have utilized the WOMAC as a primary measure. The minimal amount of MCID evidence continues to weaken research usability.

KNEE INJURY AND OSTEOARTHRITIS OUTCOME SCORE PHYSICAL FUNCTION SHORT FORM

Description

Purpose. The purpose of the KOOS-PS is to measure patients' opinions about the difficulties they experience with physical activity because of their knee problems.

Content. The KOOS-PS is a measure of physical function derived from the ADL and sport/recreation subscales of the KOOS (63). Patients rate the degree of difficulty they have experienced over the previous week due to of their knee pain with respect to the following: 1) rising from bed, 2) putting on socks/stockings, 3) rising from sitting, 4) bending to the floor, 5) twisting/pivoting on injured knee, 6) kneeling, and 7) squatting.

Number of items. The KOOS-PS contains seven items.

Response options/scale. All items are scored on a five-point Likert scale (none, mild, moderate, severe, and extreme) scored from 0 to 4.

Recall period for items. The KOOS-PS refers to the previous week.

Cost to use. Free to use. The cost of administration and data storage is unique to each practice.

How to obtain. The KOOS-PS and associated documentation are freely available at www.koos.nu.

Practical application

Method of administration. The KOOS-PS is a patient-completed questionnaire that can be completed in paper form or electronic form.

Scoring. Each question is scored from 0 to 4. The raw score is the sum of the seven items. The interval score from 0 to 100 is obtained using a conversion chart (63). There are no instructions on how to handle missing values.

Score interpretation. The possible raw score range is 0 to 28. Scores are then transformed to a score from 0 to 100, in which 0 = no difficulty.

Normative values. Not available.

Respondent burden. Based on findings for the KOOS, the KOOS-PS takes no more than 2 minutes to complete. It uses simple language and the same one-word responses for each of the seven items. Because the items relate to everyday tasks, it is not considered likely that they would have an emotional impact on the individual.

Administrative burden. The KOOS-PS takes less than 5 minutes to score using the conversion table provided (63). Training is not necessary because the questionnaire and scoring instructions are self-explanatory.

Translations/adaptations. The KOOS-PS is available in Arabic, Chinese, Danish, Dutch, English, French, German, Hindi, Italian, Korean, Norwegian, Polish, Portuguese, Russian, Spanish, Swedish, and Turkish. It can easily be compiled by extracting the seven items needed from the full KOOS forms in all languages in which the KOOS is available. Cross-cultural adaptations have been conducted for the French (64), Portuguese (65), and Turkish (66) translations.

Psychometric information

Method of development. Rasch analysis was conducted on KOOS and WOMAC data from individuals with knee OA from Sweden, Canada, France, Estonia, and the Netherlands. Patient data from 13 data sets were used (age 26-95 years; male:female ratio 1:1.4). This included community and clinical samples, such as those who had undergone previous meniscectomy, tibial osteotomy, or ACL repair, as well as those scheduled to undergo TKR (63).

Acceptability. Rates of missing data have not been reported. Findings of one study indicate no floor or ceiling effects when used in patients with knee OA (ie, no patients had the lowest or highest score, respectively) (64).

Reliability. The KOOS-PS has adequate internal consistency and test-retest reliability for groups of patients with knee OA; however, its reliability is lower than adequate for use in individuals with knee OA (Table 1). The minimal detectable change and SEM have not been reported.

Validity. *Face and content validity.* Because items are taken directly from the KOOS, which has face and content validity, this can also be assumed for the KOOS-PS, although no studies have evaluated content validity solely for KOOS-PS (58).

Construct validity. The KOOS-PS shows evidence of convergent and divergent construct validity. Higher correlations have been shown with the SF-36 physical function, role physical, and bodily pain subscales; WOMAC function subscale (excluding the KOOS-PS items); and the Osteoarthritis Knee and Hip QOL Questionnaire (OAKHQOL) physical activity domain (64,65,67). Conversely, lower correlations have been reported with the KOOS pain, symptoms, and QOL subscales; the SF-36 mental health subscales; mental health questionnaires (eg, the Profile of Mood States and the Hospital Anxiety and Depression Scale); and OAKHQOL social support (64,65,67). One study found that in patients with knee OA, the KOOS-PS had a unidimensional structure when evaluated using principal component analysis (68).

Responsiveness. In patients with knee OA, the KOOS-PS shows moderate to large effect sizes following 4 weeks of physical therapy and moderate effects 4 weeks after intra-articular hyaluronic acid injection (Table 2). The KOOS-PS is also able to discriminate between groups of patients based on use of walking aids (65). When compared directly with the WOMAC physical function subscale, the WOMAC physical function subscale was better able to detect changes over time in physical function categories (69). One study found the MCID for patients undergoing nonoperative treatment for OA to be 12 (70). For the KOOS-PS, MCID and moderate improvements were 2.2 and 15.0.

Critical appraisal of overall value to the rheumatology community

Strengths. The KOOS-PS is one of the few knee-related patient-reported outcomes that used Rasch analysis in its development. Its inclusion of only seven items facilitates its use with short measures of other dimensions, including pain VASs, and makes it ideal for those for whom long questionnaires may be onerous (eg, older populations).

Caveats and cautions. The KOOS-PS was intended for use in those with knee OA, and limited evaluation for other conditions is available. Also, using the Rasch analysis, data suggest that a 12-item short form for physical function may lead to a more optimal measurement (68).

Clinical usability. The minimal administration and scoring burden associated with the KOOS-PS make it ideal for clinical use, particularly considering that the included items are frequently asked in the standard clinical examination. However, clinicians should bear in mind that the reliability has been shown to be less than adequate for individuals.

Research usability. Psychometric testing shows the KOOS-PS to be valid and reliable for use in groups with knee OA, making it an ideal tool for measuring knee-related function in research.

KNEE OUTCOME SURVEY ACTIVITIES OF DAILY LIVING SCALE

Description

Purpose. The purpose of the KOS-ADL is to determine symptoms and functional limitation in usual daily activities caused by various knee pathologies (71).

Intended populations/conditions. The KOS-ADLS is intended for patients undergoing physical therapy for various knee pathologies, including ligament/meniscal injury, osteoarthritis (OA), and patellofemoral pain (71–73). It is applicable for patients undergoing a variety of orthopedic knee procedures and for young athletic subjects as well as older adults (74,75).

Version. Although originally described as a single index with 17 items (71), shorter versions of the KOS-ADLS have been widely used. A version using Likert-type scales is also available (76).

Content or domains. The KOS-ADLS is a single index with two sections pertaining to symptoms (pain, crepitus, stiffness, swelling, instability/slipping, buckling, and weakness) and functional limitations (difficulty walking on level surfaces, use of walking aids, limping, going up and down stairs, standing, kneeling, squatting, sitting, and rising from a sitting position) (71,76). A separate scale has been developed to assess sporting activities (71).

Number of items. The original version comprised 17 items (7 for symptoms and 10 for function), but a 14-item version (6 for symptoms and 8 for function) is also used (71,76).

Response options/scale. Patients rate items using descriptive responses, which are translated to a numerical ordinal scale for scoring. Responses for each item are scored from 0 to 5, with the exception of item 9 (0-3) and item 10 (0-2) in the 17-item questionnaire.

Recall period for items. The recall period is 1-2 days.

Cost to use. Free to use.

How to obtain. The KOS-ADLS is presented in full as an appendix in the original publication (71).

Practical application

Method of administration. The KOS-ADLS is a patient-completed questionnaire. It has not been validated for interview administration (in person or via telephone).

Scoring. The total score is calculated as the sum of scores from the responses to each item and then transformed into a percentage score by dividing by the maximum total possible score and multiplying by 100 (71,76).

Missing values. Although there are no instructions provided as to handling missing data, the original publication only analyzed questionnaires with no missing data (71).

Score interpretation. Possible transformed scores range from 0 to 100, in which 100 means no knee-related symptoms or functional limitations.

Normative values. Not available.

Respondent time to complete. It takes approximately 5 minutes to complete the KOS-ADLS questionnaire (71). No training or assistance is required because the KOS-ADLS is self-explanatory.

Administrative burden. The total score can be calculated in 5 minutes. No training is required for interpretation.

Translations/adaptations. The KOS-ADLS instrument has been validated after translation to Arabic (77,78), Chinese (79), French (80), German (81), Portuguese (82), Polish (83), Turkish (84), and Greek (85).

Psychometric information

Floor and ceiling effects. No floor effects have been detected (74,75). Acceptable ceiling effects have been reported in people with a variety of knee pathologies undergoing physical therapy and orthopedic surgeon evaluation (71,75). However, high ceiling effects have been reported 6 months after TKR (74).

Reliability. In patients with mixed knee pathologies, the KOS-ADLS has demonstrated adequate internal consistency across multiple languages as well as adequate test-retest reliability for use in groups and individuals (Table 1). A high test-retest reliability has been shown in patients with patellofemoral pain syndrome (PFPS) (86). Reliability decreases as the time increases between baseline and follow-up measurements in patients undergoing physical therapy for knee OA (87).

Validity. *Face and content validity.* During development, the KOS-ADLS was examined by orthopedic surgeons and physical therapists, who thought that it adequately covered the range of functions/painful activities performed in daily life, ensuring face validity (71). However, because item selection did not involve patient input, this instrument may lack content validity if the instruments from which items were drawn were not themselves derived from patient input (71).

Construct validity. The KOS-ADLS shows good correlation with other knee-specific scales such as the LKS (71), WOMAC subscales (74), and global assessment of function (71). Higher correlations with the physical component score than with the mental component score of the Short Form 12 indicates convergent and divergent construct validity (74).

Responsiveness. The KOS-ADLS demonstrates an ability to detect change in patients with a variety of knee disorders (Table 2). Among patients undergoing physical therapy for various knee pathologies, small effect sizes were reported at 1 week, and large effect sizes were reported at 4 and 8 weeks (71). Moderate effect sizes were reported among patients with PFPS (73). Large effect sizes have been reported following TKR (74). The responsiveness has been shown to decrease over time in patients undergoing physical therapy for knee OA (87). The PASS has not been reported.

Minimally important differences. Among patients with PFPS, the MCID has been determined to be 7.1 (73) and a minimal detectable change of 8.3% (86). In patients undergoing physical therapy for knee OA, there is an increase in the MCID from 2.2 at 2 months to 5.0 at 12 months (87).

Generalizability. The KOS-ADLS has been used in a variety of knee pathologies. It is likely generalizable to many knee conditions and many different populations because of the consistent reliability, validity, and responsiveness that is found in the literature.

Use in clinical trials. None reported.

Critical appraisal of overall value to the rheumatology community

Strengths. The KOS-ADLS is a reliable and valid instrument that is responsive to change in patients with a variety of knee conditions who are undergoing physical therapy or orthopedic procedures.

Caveats and cautions. The lack of direct patient input into item selection means that content validity cannot be assumed. The KOS-ADLS uses more descriptive responses to each item compared with other patient-reported outcomes, which may be

confusing or overwhelming for some patients, particularly those with reading difficulties. By design, the KOS-ADLS does not include items pertaining to athletic activities such as running and jumping.

Clinical usability. The KOS-ADLS is sufficiently reliable to allow use in individuals with a variety of knee disorders.

Research usability. The KOS-ADLS is reliable, valid, and appropriate for measuring change following nonsurgical and surgical interventions in a variety of knee conditions. However, researchers should be aware that if subjects being evaluated are highly physically active, this instrument is not necessarily valid. Researchers should also be consistent with which version of the scale they use.

LYSHOLM KNEE SCORING SCALE

Description

Purpose. To evaluate outcomes of knee ligament surgery, particularly symptoms of instability (88).

Intended populations/conditions. The LKS is intended for use in patients with knee ligament injury and anteromedial, anterolateral, combined anteromedial/anterolateral, posterolateral rotatory, or straight posterior instability (88).

Version. The LKS was first published in 1982 (88). The revised version (1985) added an item regarding knee locking, removed items regarding pain on giving way, swelling with giving way, and the objective measure of thigh atrophy and also removed the reference to walking, running, and jumping above the sections regarding instability, pain, and swelling (89).

Content. The original scale included the following eight items: limp; support; stair climbing; squatting; walking, running, and jumping; and thigh atrophy (88). The revised scale also includes eight items: limp, support, locking, instability, pain, swelling, stair climbing, and squatting (89).

Number of items. The LKS contains eight items.

Response options/scale. Individual items are scored differently using individual scoring scales. The revised scale modified the original scoring slightly to the following: limp (0, 3, and 5), support (0, 2, and 5), locking (0, 2, 6, 10, and 15), instability (0, 5, 10, 15, 20, and 25), pain (0, 5, 10, 15, 20, and 25), swelling (0, 2, 6, and 10), stair climbing (0, 2, 6, and 10), and squatting (0, 2, 4, and 5) (89).

Recall period for items. The recall period is not specified.

Cost of use. The revised version is freely available in the publication by Tegner and Lysholm (89).

How to obtain. The LKS can be obtained at https://stfsportsmed.com/wp-content/uploads/Lysholm-Knee-Scale.pdf.

Practical application

Method of administration. Original and revised scales were intended for in-person clinician administration (administered by the orthopedic surgeon with the patient's collaboration) (88,89), although subsequent studies have documented using the scale as a patient-completed questionnaire (90). Although significantly lower scores have been found for questionnaires versus interview administration, suggesting interview bias (91), one study reported a high level of agreement between patients and physiotherapists using a modified version of the LKS (item for swelling removed) in patients with knee chondral damage (92). Most recently, several studies have shown that telephone interviews (as opposed to face-to-face interviews) and electronically delivered questionnaires are indeed reliable modes of administration, with the perceived advantage of fostering multicenter collaborations and the potential for more accurate comparisons of outcomes between patient groups (93,94).

Scoring. Each possible response to each of the eight items has been assigned an arbitrary score on an increasing scale. The total score is the sum of each response to the eight items, of a possible score of 100. Computer scoring is not necessary.

Missing values. No instructions are provided for missing values.

Score interpretation. Possible scores range from 0 to 100, in which 100 = no symptoms or disability. Scores are categorized as excellent (95-100), good (84-94), fair (65-83), and poor (less than 64) (89).

Normative values. Normative data are available with and without stratification by sex (95,96).

Respondent burden. Time to complete has not been reported but is expected to vary depending on the administration method (ie, patient-completed versus clinician-administered methods). The LKS generally uses simple language in its questioning. However, it does use some specific medical terms, such as locking, catching, and weight bearing. Administration of this scale as it was intended (ie, clinician-administered) would ensure adequate explanation of such terms, although this may vary between clinicians. Because the items relate to everyday tasks, it is not considered that they would have an emotional impact on the individual.

Administrative burden. The LKS takes less than 5 minutes to score. Training is not necessary because the scale provides the corresponding score next to each possible response for each item.

Translations/adaptations. Since its original publication in English, several other translations have been accepted for use. An Arabic translation has been validated for OA and ligamentous and meniscal injuries (97). Cross-cultural adaptations specifically for ACL injuries have been translated and validated in the Chinese and Dutch languages (97,98). An Italian version demonstrates equivalence to the English version for assessing patellofemoral pathology (99). A German translation was found to be valid and reliable in assessing patients following total knee arthroplasty (100). The Turkish and Spanish adaptations have also been accepted for use in assessing ligamentous pathology (101).

Psychometric information

Method of development. Items pertaining to limp, support, stairs, squatting, and thigh atrophy were selected, and items for pain and swelling were adapted from the modified Larson scoring scale (102). The authors added the item for instability because they deemed this to be an important component of the disability associated with ACL injury (88). The revised scale does not report how the item for locking was selected (89). Four groups of patients with the following knee conditions were used to compare the original scale with the modified Larson scoring scale: 1) knee ligament injury and anteromedial, anterolateral, and combined anteromedial/anterolateral instability; 2) knee ligament injury and posterolateral rotatory or straight posterior instability; 3) meniscus tears; and 4) chondromalacia patellae (88). Item-response theory was not used in the development of the LKS.

Acceptability. Rates of missing data have not been reported. There are consistent reports of no floor or ceiling effects (ie, 15% of patients scoring the lowest or highest score, respectively) (75,90,103–106).

Reliability. The LKS appears to have inadequate internal consistency in patients with a variety of knee conditions (Table 1). Test-retest reliability is adequate for use in groups with knee injuries but is less than adequate for groups with mixed knee pathologies. Reliability may be inadequate for use in individuals. The minimal detectable change has been reported to be between 8.9 and 10.1 for knee injuries, whereas the SEM is reported to range from 3.2 to 3.6 for knee injuries and from 9.7 to 12.5 for mixed knee pathologies.

Validity. *Face and content validity.* The LKS has been reported as having face validity, as evaluated by five orthopedic surgeons with sports medicine experience (75). Because the items in the LKS are surgeon derived, content validity from the patient's perspective cannot be assumed.

Construct validity. Multiple studies have reported convergent construct validity for the Lysholm score, finding significant correlations with the Hospital for Special Surgery modified knee ligament rating system, Cincinnati Knee Ligament Score, the IKDC Subjective Knee Evaluation Form, the Fulkerson and Kujala scores, and the WOMAC (105–107). Two studies have reported evidence of convergent and divergent construct validity, finding the Lysholm score to correlate more highly with the physical components of the Short Form 12 and SF-36 than with the mental components (75,90). The Lysholm score was shown to satisfy the Rasch model after the removal of the item for swelling in patients awaiting surgery for knee chondral damage (92).

Ability to detect change. Large effect sizes have been reported following ACL reconstruction (6-9 months postoperative), meniscal repair (1 year postoperative), and microfracture (1-6 years postoperative) (Table 2). Large effect sizes are also reported following 1 month of physical therapy in a group of patients with mixed knee pathologies.

Minimal important difference. The MCID and PASS have not been calculated in any patient population. Specifically, when comparing responsiveness following autologous chondrocyte implantation, the Lysholm and IKDC Subjective Knee Evaluation Form were the most sensitive to detecting changes compared with the Modified Cincinnati Knee Rating System, KOOS, and SF-36 (108).

Critical appraisal of overall value to the rheumatology community

Strengths. The LKS is a freely available measure that is able to detect change following nonsurgical and surgical intervention. It is considered to have face validity by orthopedic surgeons. Because the LKS assesses everyday activities as opposed to higher functional activity, delayed return to sport has little impact on the LKS. Therefore, the LKS may be ideal for assessing short-term outcomes or outcomes in patients not intending to return to a specific sport (108).

Caveats and cautions. Content validity cannot be assumed because the items included in the LKS were surgeon derived. The LKS was developed as a clinician-administered tool, which increases the potential for interviewer bias if the patient-reported outcome is applied as intended. Despite this, there are inconsistencies between methods of administration of the LKS in published studies. The MCID and PASS are lacking in psychometric analysis.

Clinical usability. Minimal administrative and respondent burden makes the LKS attractive for clinical use. The lack of floor and ceiling effects across different knee conditions suggests that the LKS is useful for tracking improvement with intervention as well as deterioration over time in patients with various knee pathologies. However, clinicians should consider the impact of inadequate reliability in evaluation of individuals.

Research usability. The LKS is reliable for use in research on ligamentous injuries of the knee, chondral injuries, and patellar dislocation. The use of the LKS and IKDC Subjective Knee Evaluation Form together has proven to represent a responsive combination for efficiently evaluating treatment effects following autologous chondrocyte implantation (108). It is important that researchers consistently use the same scale version (89). Researchers should be aware that the psychometric properties may change between different administration methods, ensure consistent administration within and between studies, and be aware that clinician and patient ratings may differ substantially. The lack of a known MCID is a weakness.

TEGNER ACTIVITY SCORE

Description

Purpose. To provide a standardized method of grading work and sporting activities (89). Developed to complement the LKS based on observations that limitations in function scores (Lysholm) may be masked by a decrease in activity level (89).

Intended populations/conditions. The TAS is intended for use in conjunction with the LKS, originally in patients with ACL injury (89).

Version. Although in some circumstances it has been modified slightly to accommodate different populations, the standard TAS remains in its original format.

Content. The TAS consists of a graduated list of ADLs, recreation, and competitive sports. Patients select the level of participation that best describes their current level of activity.

Number of items. One item is selected from a list of 11.

Response options/scale. A score of 10 is assigned based on the level of activity that the patient selects. A score of 0 represents sick leave or disability pension because of knee problems, whereas a score of 10 corresponds with participation in national and international elite competitive sports (89). Activity levels 6 to 10 can only be achieved if the person participates in recreational or competitive sport.

Recall period for items. The TAS refers to current ability.

Cost to use. Freely available in the original publication (89).

Practical application

Method of administration. The TAS was originally established as an in-person, clinician-administered tool (109) but has been used more recently as a patient-completed questionnaire (90,110).

Scoring. A score ranging from 0 to 10 is assigned based on the level of activity that the patient selects as best representing their current activity level. Computer scoring is not necessary.

Missing values. Not applicable (single score).

Score interpretation. Possible scores range from 0 to 10. Higher scores represent participation in higher-level activities.

Normative values. Normative data have been presented by sex and age group (95).

Respondent burden. The TAS is reported to take a mean (SD) of 3.3 (0.6) minutes to complete for those who have undergone TKR (111). The scale classifies work, recreational, and sport activities in a graded activity scale using common terminology. As such, patients should not have difficulty selecting which level corresponds with their current activity. Degree of difficulty (measured on a VAS) has been reported to increase with age ($r = 0.25$; $P = 0.03$) (111).

Administrative burden. Scoring time is negligible because the score is based on a single selected item. Training is not necessary.

Translations/adaptations. The TAS is available in English. Cross-cultural translations are now validated for use in Swedish, Dutch, German, Chinese, and Iranian populations with ACL injuries, and a German translation has been validated for use in the total knee arthroplasty population (98,100,112–114). Use in other rheumatology populations has consisted of ankle and shoulder disorders.

Psychometric information

Method of development. Orthopedic surgeons selected items they believed to be difficult for patients with ACL injury. Forty-three patients with ACL-deficient knees then completed a questionnaire in which they graded these activities according to how difficult they were. This formed the basis of item selection for the TAS. Both paper and electronic forms have been found to be reliable methods of administration (93).

Acceptability. Studies consistently report no floor or ceiling effects in those with knee injury or OA (ie, 15% scored the lowest or highest score, respectively) (90,100,103,111).

Reliability. The TAS has adequate test-retest reliability for groups with knee injuries and knee OA, although reliability is less than adequate for use in individuals (Table 1). For knee injuries, the minimal detectable change is 1, whereas the SEM ranges from 0.4 to 0.64.

Validity. *Face and content validity.* At face value, the TAS covers a wide variety of activity levels that may be applicable to patients with ACL injuries and other knee injuries. However,

because the initial activity selection was conducted by orthopedic surgeons with patient input afterward regarding the difficulty of these selected activities, content validity cannot necessarily be assumed.

Construct validity. Evidence for convergent and divergent construct validity is provided by studies that found higher correlations with the physical component of the Short Form 12 than the mental component (90,100,111). The TAS has also shown significant correlations with the IKDC Subjective Knee Evaluation Form, the Knee Society Score function score, the WOMAC pain and function subscales, and the OKS (90,100,103,111).

Generalizability. The TAS was found to be reliable in both adult and pediatric populations (115).

Ability to detect change. Following meniscal surgery, moderate effect sizes are seen 12 months postoperatively in those with isolated meniscal lesions, and large effect sizes are seen in those with combined lesions (Table 2). In those who have undergone ACL reconstruction, effect sizes are reported to be moderate at 6 months and large at 9 months, 1 year, and 2 years. The MCID and PASS have not been determined.

Critical appraisal of overall value to the rheumatology community

Strengths. The TAS is a simple, freely available measure of activity level that spans work, sporting, and recreational activities. It is one of the few PROMs that were developed to consider the influence of activity level on other symptoms, including pain alleviation when aggravating activities are avoided. The TAS was found to have clear benefits over other PROMs and is the preferred PROM for ACL injuries in the United Kingdom (116).

Caveats and cautions. The TAS was originally intended and developed for patients with ACL injury as an adjunct to the LKS and not as a stand-alone measure. The MCID is missing from psychometric analysis. Studies suggest that TAS data need to be adjusted for age and sex (117).

Clinical usability. Clinicians should note that the reliability of the TAS may be inadequate for use in individuals.

Research usability. Although valid and reliable for use in groups, use of the TAS in research may need to be applied with caution. Given its intent to measure change within patients, the TAS may be more appropriate for within-subject repeated-measures studies rather than between-group comparisons.

OXFORD KNEE SCORE

Description

Purpose. The OKS is a brief questionnaire for patients undergoing TKR that reflects the patient's assessment of their knee-related health status and benefits of treatment (118).

Intended populations/conditions. The OKS is intended for patients undergoing TKR.

Version. A new version was proposed on the basis that some surgeons believed that the scoring of the original version was nonintuitive (ie, lower scores represented better outcomes and higher scores represented worse outcomes), in which the original 12 items are used but the scoring is different (119).

Content or domains. The OKS is a single index pertaining to knee pain and function (pain severity, mobility, limping, stairs, standing after sitting, kneeling, giving way, sleep, personal hygiene, housework, shopping, and transport). The questionnaire can be separated into pain and function subscales with good validity and responsiveness (120).

Number of items. The OKS contains 12 items.

Response options/scale. Each item is followed by five responses (scores range 1-5), in which 1 is the best outcome and 5 is the worst outcome. The modified version also has five responses to each item, but the scoring is from 0 to 4, in which 0 is the worst outcome and 4 is the best outcome.

Recall period for items. The recall period for the OKS is the previous 4 weeks.

Cost to use. Free to use.

How to obtain. The original version can be found in its original publication (118). The modified version is freely available online (http://www.orthopaedicscore.com/scorepages/oxford_knee_score.html) (119).

Practical application

Method of administration. The OKS is a patient-completed questionnaire.

Scoring. Originally, each response to each item was assigned a score from 1 to 5 (1 = no problem and 5 = significant disability). The modified version assigns a score from 0 to 4 (4 = no problem and 0 = significant disability). The total score is calculated as the sum of scores from the responses to all 12 items.

Missing values. No instructions are provided for missing values.

Score interpretation. In the original version, the total score ranges from 12 to 60 (118), whereas in the modified version, the total score ranges from 0 to 48 (119). Higher scores in the original version reflect worse outcomes, and lower scores reflect better outcomes. In the modified version, this is reversed.

Normative values. Not available.

Respondent time to complete. The OKS is reported to involve minimal respondent burden (118). It takes approximately 5 to 10 minutes to complete the questionnaire. No training or assistance is required because the questions are self-explanatory.

Administrative burden. Scoring is simple and quick (118). Calculation of the total score takes 1 to 5 minutes. No training is necessary.

Translations/adaptations. The OKS is translated and validated in many languages, including Arabic (7,121), Chinese (122–124), Finnish (125), German (126), Japanese (127), Korean (128), Persian (129,130), Portuguese (131), Spanish (132), Swedish (133), Thai (134), and Turkish (135).

Psychometric information

Floor and ceiling effects. A study reported no floor or ceiling effects prior to TKR (74). Six months postoperatively, although there were no floor effects, there were ceiling effects reported (27% of patients scored the top score). Conversely, one study found large floor effects at an average of 18 months postoperatively, with no ceiling effect (136). Furthermore, the ceiling effect has been shown to increase from 6 months to 2 years postoperatively from TKR, whereas no floor effect was found at either time point (137). Another study found no floor or ceiling effects in patients with knee OA (132).

Reliability. The OKS has adequate internal consistency across multiple languages (118,122,126,127,133,134) (Table 1). The original study reported adequate test-retest reliability for use in groups and individuals (118). The test-retest reliability was confirmed in patients with OA being managed nonoperatively (138).

Validity. *Face and content validity.* Extensive input from patients in the development of the OKS ensures content validity.

Construct validity. The OKS shows good correlation with knee-specific and general health questionnaires such as the WOMAC, the American Knee Society Score, KOS-ADL, and the SF-36 and Health Assessment Questionnaire pain and physical function components (118,139). Internal and external validity is adequate postoperatively (136). Convergent and divergent construct validity is demonstrated by higher correlations with the physical component of the Short Form 12 than with the mental component (74). Convergent and divergent validity have also been confirmed when the pain and function subscales of the OKS has been separately compared with other outcome scores (138). The OKS has been shown to fit Rasch models following the rescoring of some items (140) and the removal of items for limp and kneeling (141).

Responsiveness. The OKS demonstrates good sensitivity and responsiveness to change (Table 2). Responsiveness is consistent using raw OKS data or after Rasch analysis (137). Large effect sizes have been reported 6 to 12 months after TKR (118,142) and 1 year after high tibial osteotomy (143). The OKS has also been found to be a good predictor of revision TKR within 6 months (144) and is also a predictor of range of motion after TKR (145). The effect size is larger in patients who report positive changes in their knee symptoms over time compared with patients who report a negative progression in symptoms (132).

Minimally important differences. The minimum detectable change with a 95% confidence interval after high tibial osteotomy was reported to be 8.29 (143). The minimum detectable change with a 90% confidence interval (MDC90) and MIC 6 months after TKR were found to be 4.15 and 9.22, respectively (146). In patients with knee OA, the MDC90 after 3 months of nonoperative management was 6, and the minimum important difference was 6.4 (138). In this same patient cohort, the MIC for the total OKS was 7.1, the MIC for OKS the pain subscale was 17.3, and the MIC for the function subscale was 10.6. The PASS has not been reported.

Generalizability. The OKS was intended for use in patients with knee OA before and after TKR. It is likely generalizable to many knee conditions and many different populations because of the consistent reliability, validity, and responsiveness that is found in the literature.

Use in clinical trials. None reported.

Other. Patients are able to recall their preoperative health status regarding the OKS with good consistency (147) even up to 1 year postoperatively (148).

Critical appraisal of overall value to the rheumatology community

Strengths. The OKS is a self-administered questionnaire developed to measure outcomes following TKR. Because of its simplicity and ease of administration it has been used widely, especially in the United Kingdom, and is available in languages other than English. For the same reasons, it can be used as a cost-effective screening tool in short-term (2 years) follow-up of TKR compared with physician-administered instruments such as

the American Knee Society Score, as reported by one study (149). It may be separated into pain and function subscales.

Caveats and cautions. Although simple, some items are double barreled and may be confusing to patients (eg, trouble getting in and out of a car or using public transportation). Some response options potentially overlap with others, which may also cause confusion. The use of an aggregate score combining pain and function may mask changes in one domain, particularly given that only one of the 12 items relates solely to pain. If the OKS is separated into pain and function subscales, the administrator must be aware of the complex interaction between pain and function and therefore patient interpretation of the questions. The floor and ceiling effects postoperatively make postoperative comparisons and distinctions more difficult.

Clinical usability. Psychometric testing suggests that the OKS is sufficiently reliable for use in individuals with knee OA. The ease of administration and scoring makes it a useful tool for clinical use. However, clinicians should be aware that some patients may require explanation of individual items, which could introduce interviewer bias.

Research usability. The OKS is a knee OA–specific measure that is reliable, valid, and responsive to change following TKR. Researchers should be aware of the different scoring methods when interpreting findings of previous research. It is correlated with the American Knee Society Score, and therefore these scores can be directly compared (150).

WESTERN ONTARIO AND MCMASTER UNIVERSITIES OSTEOARTHRITIS INDEX

Description

Purpose. To assess the course of disease or response to treatment in patients with knee or hip OA (151,152).

Intended populations/conditions. The WOMAC is intended to be used in patients with knee and hip OA (151,152).

Version. Initially developed in 1982, the WOMAC has undergone multiple revisions (most recent version 3.1). It is available in a five-point Likert scale, a 100-mm VAS, and an 11-box numerical rating scale (153,154). Reduced and modified versions of the WOMAC have been validated but are not endorsed on the WOMAC website (155–158).

Content or domains. The WOMAC contains the following three subscales: 1) pain severity during various positions or movements, 2) severity of joint stiffness, and 3) difficulty performing daily functional activities.

Number of items. The WOMAC consists of 24 items.

Response options/scale. In the Likert version, each item offers five responses: none = 0, mild = 1, moderate = 2, severe = 3, and extreme = 4. Alternatively, the VAS and numerical rating scale versions permit responses to be selected on a 100-mm or 11-box horizontal scale, respectively, with the left end marked as none and the right end marked as extreme (151,152).

Recall period for items. The WOMAC has a recall period of 48 hours.

Cost to use. The WOMAC is not free; the cost depends on the research project.

How to obtain. The WOMAC is available from Professor Nicholas Bellamy (n.bellamy@uq.edu.au). To obtain licensing and fee information and permission to use the WOMAC for clinical or research purposes, a request needs to be submitted at http://www.womac.com.

Practical application

Method of administration. The WOMAC is a self-administered or interview-administered questionnaire. It has been validated for use in person, over the telephone, or electronically via a computer or mobile phone (152,159–162). Electronic and paper questionnaires show high agreement (163).

Scoring. The total score for each subscale is the sum of scores for each response to each item and can be calculated manually or using a computer. The ranges for possible subscale scores in the Likert format are as follows: 0 to 20 for pain (5 items each scored 0-4), 0 to 8 for stiffness (2 items), and 0 to 68 for physical function (17 items). In the VAS format, the ranges for the three subscale scores are 0 to 500 for pain, 0 to 200 for stiffness, and 0 to 1700 for physical function (151,152).

Missing values. If two or more pain items, both stiffness items, and four or more physical function items are missing, the response should be regarded as invalid and the deficient subscale(s) should not be used in analysis (151).

Score interpretation. Higher scores indicate worse pain, stiffness, or physical function.

Normative values. Australian population-based normative data stratified by age and sex have been reported (164).

Respondent time to complete. The WOMAC takes 5 to 10 minutes to complete.

Administrative burden. The WOMAC takes approximately 5 minutes to score. Training is not necessary.

Translations/adaptations. The WOMAC version 3.1 is available in over 100 languages (153) and has validated language translations for Arabic (165), reduced Arabic (121), Bangladesh (166), Chinese (167,168), Finnish (169), German (170), Greek (171), Hebrew (172), Italian (173), Japanese (174), Korean (175), Moroccan (176), Nepali (177), Persian (129), Brazilian Portuguese (178), Singapore (179), Spanish (180), Swedish (181,182), Thai (183), and Turkish (184,185).

Psychometric information

Floor and ceiling effects. Reports of floor and ceiling effects have differed between studies (74,167,183,185,186). The stiffness subscale has been reported as having floor and ceiling effects prior to intervention (74,167,185), as well as up to 1 year postoperatively from TKR (187). Ceiling effects have been reported by various studies for all subscales 6 months and 2 years after TKR (74,186).

Reliability. The stiffness and function subscales have consistently demonstrated adequate internal consistency in knee OA (Table 1). Studies have generally reported adequate internal consistency for the pain subscale, although there have been reports of slightly lower than adequate internal consistency. There have been mixed findings regarding adequacy of test-retest reliability in knee OA for all subscales. Test-retest reliability for the stiffness subscale may not be adequate for use in individuals with knee OA. One study that investigated test-retest reliability in patients with chondral defects found that all subscales had adequate reliability for use in groups but that only the function subscale was adequate for individual use. However, the putting on socks item of the physical function subscale may present problems in stability and variance when analyzed with the Rasch model (188). Additionally, the physical function subscale may have a stronger association with pain than performance (189,190). The minimal detectable change and SEM vary according to condition and subscale. These measures tend to increase over time, and the reliability decreases over time (87).

Validity. *Face and content validity.* Because the WOMAC was developed with extensive input from patients with OA as well as input from academic rheumatologists and epidemiologists experienced in the clinical assessment of rheumatologic diseases, the WOMAC can be considered to have face and content validity.

Construct validity. Multiple studies have shown that the WOMAC subscales demonstrate good construct validity. Moderate to strong correlations with measures of similar constructs (eg, the SF-36 physical subscales, pain/handicap VAS) suggest convergent construct validity (167,172,173,176,184,185,191, 192), whereas lower correlations with measures such as the SF-36 mental subscales indicate divergent construct validity (167,173,184,185,192). Convergent validity was also demonstrated with strong correlation with the 30-second chair stand and 50-foot timed walk tests (193). Although Rasch analyses have largely utilized mixed knee and hip OA cohorts, it has been reported that there is no differential item functioning based on affected joint (194). Although one study found the pain subscale to demonstrate good item separation and unidimensionality in patients with knee or hip OA (195), a subsequent study found that a reduced pain subscale (with night pain and pain on standing removed) fit the Rasch model and provided more stable results over time and between patients with knee or hip OA and those who have undergone joint replacement (194). The function subscale demonstrates more variability. Although found to have good item separation and unidimensionality in knee/hip OA, function items for performing light chores, getting in/out of a car, and rising from bed were found to be redundant (195). Similarly, Davis et al (194) suggested a 14-item function subscale, with items for heavy domestic duties, getting in/out of the bath, and getting on/off the toilet removed. There is a strong correlation with psychological measures and total WOMAC score, indicating poor divergent validity (196). Hip abductor and knee extensor strength are not correlated with WOMAC function subscale (197).

Responsiveness. The WOMAC appears to be responsive to change following surgical and nonsurgical interventions for knee OA and chondral defects (Table 2) (198). A recent study has confirmed the high responsiveness in patients undergoing TKR, with a mean change in score of 29 at 3 months postoperatively (199). In particular, the physical function domain has been suggested to be the best choice for detecting changes over time compared with other measures (69). However, in patients undergoing exercise therapy, the total WOMAC was found to be more responsive than the physical function subscale (200). In patients with knee OA, large effect sizes are consistently reported on all three subscales up to 2 years post-TKR. This was recently confirmed at 1 year postoperatively; however, the stiffness subscale has a smaller effect size than the pain or function subscales (201). Furthermore, effect size decreases over time up to 2 years after TKR (187). Following exercise intervention, the stiffness subscale shows small effect sizes at 2 weeks compared with moderate to large effect sizes for the pain and function subscales; however, these also are small at 6 months. Acupuncture has shown small to moderate effect sizes in the short term (3 weeks) but large effect sizes after 8 weeks. Drug intervention tends to show different patterns across 12 weeks for the three subscales.

Effect sizes for pain tend to be large initially (1 week) and become more variable at 6 weeks (moderate to large) and 3 months (small to large). In comparison, the stiffness subscale tends to show small to moderate effect sizes over the initial 4 weeks, becoming moderate to large by 3 months. Similarly, the effect sizes for function also gradually increase, starting at moderate at 2 weeks

and becoming moderate to large at 6 and 12 weeks. Following surgery for chondral defects, large effect sizes are seen for pain and function at 6 and 12 months postoperatively, whereas moderate effect sizes are seen on the stiffness subscale. Using composite WOMAC outcomes by combining the subscales improves responsiveness and reduces the necessary sample size (202). The reduced WOMAC has been shown to have similar responsiveness to the original WOMAC (203).

Minimally important differences. The MCID has been calculated for TKR (up to 2 years postoperatively; range for pain 22.9-36, range for symptoms 14.4-21.4, and range for function 19-33) and nonsteroidal antiinflammatory use (4 weeks; function 9.1). At 3 months postoperatively, the MCID was determined to be 10.21 (199). The PASS has been determined to be 31.0 (95% confidence interval 29.4-32.9) for the function subscale in people with knee OA (204). The MIC has also been determined for the short WOMAC as 7.9 and 9.8 points for small change, 8.4 and 9.8 points for medium change, and 12.1 and 10.1 points for large change (205).

Generalizability. The WOMAC has been mainly used for OA and TKR; however, it has been used in other knee pathologies. Because of the consistent reliability, validity, and responsiveness that is found in the literature, it can be inferred that the WOMAC is generalizable to many knee conditions and many different populations.

Use in clinical trials. The WOMAC was used to assess efficacy of the addition of oxygen therapy to usual therapy in patients experiencing a flare of knee arthritis (206). The effect of patellofemoral overstuffing on clinical outcomes was also investigated with the WOMAC as the primary outcome (207). The WOMAC was a primary outcome measure in a trial investigating tanezumab for hip and knee arthritis (208). It has also been used as the primary outcome measure in a trial assessing the impact of change in physical activity on pain and physical function (209). A trial of aqueous extract of *Terminalia chebula* fruit as a dietary supplement in healthy adults who were overweight used the modified WOMAC as a primary outcome (210). Pain progression evaluated using the WOMAC was associated with radiographic and magnetic resonance imaging evaluation of cartilage loss in symptomatic knee osteoarthritis (211).

Critical appraisal of overall value to the rheumatology community

Strengths. The WOMAC is one of the most commonly used patient-reported outcomes for knee OA. It is simple and quick to administer and score using guidelines provided. The utilization of patients in development ensures content validity. In addition, the WOMAC has undergone validated translations into multiple languages. The use of individual scores for each subscale, rather than an aggregate score, enhances interpretation.

Caveats and cautions. The need to obtain permission and pay licensing fees prior to use may encourage researchers and clinicians to seek alternatives. The inclusion of tasks in the function subscale that may not be performed regularly by all patients (eg, stair climbing and taking a bath) may result in missing data. Content validity is not ensured for more physically active patients because the function scale does not include more difficult functional tasks. Rasch analysis suggests that the function subscale contains redundant items. The physical function domain has a stronger association with pain than performance. Patients have to recall symptoms during specific movements. There is a correlation between patient psychological status and WOMAC score. Reliability and responsiveness decrease with time.

Clinical usability. The variability in administration methods makes the WOMAC a good choice for clinical use, particularly when dealing with patients with communication difficulties. Minimal floor effects mean that the pain and function subscales are able to monitor deterioration in condition over time, whereas ceiling effects have only been reported following TKR. However, clinicians should consider that the stiffness subscale may not be sufficiently reliable for use in individuals. An additional physical function measure may be employed to ensure that this construct is fully measured because of its association with pain.

Research usability. Psychometric testing indicates that the WOMAC is sufficiently reliable and valid for use in research. The variety of validated language translations and methods of administration is a major strength for WOMAC use in research. A body of research supports the responsiveness to change of the WOMAC following surgical and nonsurgical interventions. The extensive use of the WOMAC in previous research facilitates the comparison of new findings.

ACTIVITY RATING SCALE

Description

Purpose. The ARS was developed as a short, simple knee-specific questionnaire to evaluate the activity level of patients with various knee disorders who participate in different sports. It is intended to provide data on an athlete's highest activity level within the past year (ie, at a time when they were most active) (212).

Intended populations/conditions. The ARS is intended for use in various knee conditions, including ligament, meniscus, and chondral injury; patellofemoral pain; osteochondritis dissecans; trabecular fracture; and iliotibial band syndrome (212).

Version. There have been no modifications to the original version.

Content. The ARS is a single index pertaining to frequency of the following athletic activities: running, cutting, decelerating, and pivoting.

Number of items. The ARS consists of four items.

Response options/scale. Each item is followed by five responses for the frequency of each functional component within the past year.

Recall period for items. The recall period for the ARS is 1 year.

Endorsements. None.

Examples of use. The ARS has been used in various conditions, including ACL injury, cartilage injury, and knee OA, and interventions, including ACL reconstruction, autologous chondrocyte implantation, microfracture, high tibial osteotomy, and TKR.

Practical application

How to obtain. The ARS can be found as an appendix in the original publication (212).

Method of administration. The ARS is a patient-completed questionnaire, administered either on paper or electronically, with particularly high rates of agreement (213). The ARS not yet been validated for interview administration (by telephone or in person).

Scoring. Each item is scored from 0 to 4, in which 0 = less than one time a month, 1 = one time in a month, 2 = one time in a week, 3 = two to three times in a week, and 4 = four or more times in a week. The total score is the sum of scores from responses to each of the four items (212).

Missing values. There are no specific instructions for handling missing values.

Score interpretation. The total possible score range is 0 to 16, in which 16 represents more frequent participation.

Normative values. Not available.

Respondent burden. The ARS takes approximately 1 minute to complete. The respondent burden was intentionally minimized through the inclusion of only four items (212).

Administrative burden. The ARS takes less than 5 minutes to score. No training is required.

Translations/adaptations. A cross-cultural adaptation has been conducted for the Swedish translation (214), and a Persian version has been translated and validated specifically for ACL injuries (215).

Psychometric information

Method of development. Items were selected by literature review and expert opinion (orthopedic surgeons who specialized in sports medicine, physical therapists, and athletic trainers) and by surveying patients with knee disorders. Item reduction involved 50 patients with a variety of knee disorders who were physically active who rated the importance and difficulty associated with each functional task on the preliminary list. The top four, as agreed by the panel of clinicians, were retained in the final version (212).

Acceptability. Information on missing data and floor/ceiling effects is not available.

Reliability. One study has evaluated the test-retest reliability of the ARS, finding adequate reliability for use in groups and individuals (212) (Table 1). The internal consistency has not been reported.

Validity. *Face and content validity.* The use of patients with knee disorders in both item selection and reduction ensures content validity. The final item selection also involved the opinion of clinicians to ensure face validity (212).

Construct validity. The ARS has been reported to have moderate to strong correlation with other knee-related scales that measure activity levels, such as the TAS, the Cincinnati Knee Ligament Score, and the Daniel Score, suggesting good convergent construct validity (212).

Generalizability. Previously, the ARS had only been validated for adult use. However, a study published in 2015 found the ARS to be reliable in patients younger than the age of 18 with knee injuries, with decreasing reliability in patients younger than 14. Test-retest data confirmed its reliability in all but one of the questions in the cohort with participants aged between 14 and 18. Although the questionnaire may prove useful in this pediatric population, its usefulness may be limited by the significant ceiling effect observed; more than half of the patients had maximum scores of 16 (50.6%) (216).

Ability to detect change. The responsiveness, MCID, and PASS have not been reported (Table 2). Rasch analysis was not performed.

Critical appraisal of overall value to the rheumatology community

Strengths. The ARS is a short, simple measure that represents minimal administrator or respondent burden. Because it assesses four common components of various sporting activities, rather than nominating specific sports, it is generalizable across a wide range of elite and recreational athletes. In addition, to the

extent that activities such as running, stopping, and changing direction are also needed for nonsport activities, it could be applicable to other situations (eg, work tasks).

Caveats and cautions. Because its focus is limited to specific activities, it is important to assess activity-related scales in conjunction with questionnaires used to evaluate pain and function, because activity level may be particularly important as a potential confounding variable when evaluating patient outcomes following knee injuries. Often, an inverse relationship is observed when administered together. Some patients may report pain and functional limitations but are able to return to a higher level of activity. On the other hand, perhaps the higher level of activity is associated with increased pain and perceived limitations. Inversely, patients may report better outcomes in pain and function but report lower ARS scores as a result of lifestyle changes made to avoid symptoms and risk of reinjury (217). Therefore, the utility of the ARS is maximized as an adjunct to scales that assess other domains of knee function (218).

Other activities such as swimming and jumping cannot be evaluated by this scale. Furthermore, because the ARS does not focus on current ability but on baseline activity frequency possibly prior to injury, the validity of the instrument depends on the patient's accurate recollection of this frequency. The accuracy of such recollection may be influenced by the time since injury and by the current state of activity. Lack of evidence for responsiveness to change/sensitivity is also a limitation. The ARS should be used as an adjunct to other knee instruments assessing symptoms and difficulty (212).

Clinical usability. The ARS is a short activity–specific questionnaire, making it good for clinical use. It would be suitable for patients who participate in land-based sports or activities that do not involve jumping as a primary movement. Clinicians should consider that the 1-year recall period may be difficult for some patients.

Research usability. The lack of psychometric data for the ARS limits its use in research. Because the scale measures the highest level of activity over the past year without taking time of injury into account, it may be more suited for within-subject study designs rather than for comparing ratings between subjects.

PATIENT-REPORTED OUTCOME MEASUREMENT INFORMATION SYSTEMS PHYSICAL FUNCTION

Description

Purpose. The PROMIS-PF was developed to measure self-reported capability rather than actual performance of physical activities. This includes the functioning of one's upper extremities (dexterity), lower extremities (walking or mobility), and central regions (neck and back) as well as instrumental ADLS such as running errands (219).

Content. The PROMIS, funded by the National Institutes of Health, was developed to be a tool for both clinicians and researchers to access efficient, precise, valid, and responsive adult and pediatric PROMs in health and well-being (220). This tool is unique because it is useful in various disciplines in measuring physical, mental, and social health in individuals with chronic conditions (219). There are multiple subscales specific to the goals of measurement and patient population. The physical function form of PROMIS specifically measures the ability to carry out various activities that require capability, ranging from self-care to more vigorous activities of mobility, strength, and endurance (219).

Number of items. Form 10a has 10 items. The first five focus on the degree to which the patient's health limits the following activities: vigorous walking, climbing stairs, carrying groceries, and bending or kneeling. The second five focus on difficulty in carrying out the following ADLs: vacuuming or yard work, dressing, shampooing hair, washing and drying the body, and using the toilet.

Recall period for items. There is no specification of a recall period.

Cost to use. The PROMIS forms are free to use in the single-use forms. Integrated data collection and computerized scoring are priced independently and will vary based on chosen system and needs.

How to obtain. The PROMIS-PF is available at https://www.assessmentcenter.net/PromisForms.aspx (http://www.healthmeasures.net/index.php?option=com_content&view=category&layout=blog&id=71&Itemid=817).

Practical application

Method of administration. PDF forms as well as integrated data collection tools are available through HealthMeasures (http://www.healthmeasures.net/resource-center/data-collection-tools).

Scoring. Creators of the PROMIS intended the measurement to be scored according to response pattern scoring, with item-level calibrations using the HealthMeasures scoring service (https://www.assessmentcenter.net/ac_scoringservice). However, there is also a table to be used. Each question has five response options (Likert) ranging from 1 to 5. A score of 5 is equivalent to no limitation or difficulty, and a score of 1 is equivalent to being unable to complete.

Scoring interpretation. The raw score ranges from 10 to 50, and the scaled score ranges from 13.5 to 61.9, with 50 or 61.9 representing optimal physical function. The T-score rescales the raw score into a standardized score with a mean of 50 and an SD of 10. Therefore, a person with a T-score of 40 is 1 SD below the mean.

Respondent time to complete. The PROMIS-PF takes 5 minutes to complete.

Administrative burden. Multiple integrative data options for all the PROMIS measures exist that would alleviate any significant administrative burden, but these come with variable price points. However, it takes about 3 to 5 minutes to score manually on the single-use PDFs.

Translations and adaptations. The PROMIS-PF 10a is available in the following languages: English, Spanish, Danish, Dutch, French, German, Hungarian, Italian, Polish, Russian, simplified Chinese (Mandarin), traditional Chinese, and Ukranian.

Psychometric information

Floor and ceiling effects. Studies have shown no floor or ceiling effects with meniscal injuries, patellofemoral malalignment, multiligamentous injuries, and chondral disease (221–223). In a study of 204 patients, the PROMIS-PF was found to have no floor effect, although one patient scored the highest possible score (224). When compared with the KOOS-ADL, KOOS sport subscale, and SF-36 physical function subscale, the PROMIS-PF had the lowest ceiling effect of the instruments, with 1.4% at 6 months and 9.0% at 2 years in patients with ACL injuries, which is well below the 15% cutoff (225).

Reliability. Few studies have demonstrated reliability in the PROMIS-PF. One study showed high reliability in patients with rheumatoid arthritis (226). Hung et al, when attempting to validate the lower extremity physical function computer adaptive test based on PROMIS-PF items, found the items to demonstrate high reliability (227).

Validity. Hung et al, when attempting to validate the lower extremity physical function computer adaptive test based on PROMIS-PF items, found the item bank to be unidimensional and free of item bias, with high content and construct validity (227). Another study by the same group showed adequate face validity as well as construct validity (228). Good construct validity of the PROMIS-PF in patients with rheumatoid arthritis has also been shown (226,229). Content validity was further shown in patients with tenosynovial giant cell tumors of the knee (230). Strong validity of comparisons for the PROMIS-PF items was shown in patients with different musculoskeletal disorders, namely, chronic pain, rheumatoid arthritis, and OA, and there is high correlation with SF-36 scores (224,231). Good convergence has been found between the PROMIS-PF, KOOS, and IKDC scales (232).

Responsiveness. The PROMIS assessments collectively have been shown to be very responsive to change (233). When compared with the KOOS-ADL, KOOS sport subscale, and SF-36 physical function subscale, the PROMIS-PF showed equal responsiveness and excellent utility in the postoperative ACL course (225). The PROMIS-PF was specifically found to have high responsiveness to patients with OA, whereas pain, depression, and anxiety PROMIS forms only have moderate responsiveness (234). It has also been shown to compare well with disease-specific scales in regard to knee arthroscopy patients (232).

Critical appraisal of overall value to the rheumatology community

Strengths. The PROMIS-PF can be used not only for a variety of conditions in the knee but also for a variety of musculoskeletal and rheumatological conditions. Furthermore, with regard to the knee, it is comparable in psychometric properties with disease-specific scales. There is low burden to the patient as well as low administrative burden. There are also a lot of resources available to integrate scoring and maintaining data.

Caveats and cautions. PROMIS has multiple subscales and forms that can be used. It is important to use the best subscale for a given need. Psychometric properties have not been assessed for all knee conditions, and the PROMIS-PF score was not developed specifically for knee conditions.

Clinical usability. The PROMIS-PF is easy to use and has low respondent burden and administrative burden. It can be used for many conditions.

Research usability. The PROMIS-PF is easy to use in the research setting and has been shown to be comparable with other scales that are specific to knee conditions. It may allow for comparing similarities in physical function changes between patients of varying conditions.

CONCLUSIONS

We reviewed nine of the instruments that have been developed to measure patient-reported knee function and one measure that has been used for overall physical function but has been adequately tested for assessment of knee-related conditions. Since the last review of some of these knee measures was published in 2011, there has been an enormous body of research evaluating their psychometric properties in patients with varying knee conditions. Furthermore, many tools have also been cross-

culturally translated into multiple languages and adapted when needed. Although other measures may be useful (235), this extensive review provides researchers with the necessary information for the nine most commonly used instruments in trials in the last 10 years as well as information for the PROMIS-PF. When seeking to use knee measures, it might be useful to refer to core outcome sets or minimum standard sets of outcomes (236).

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. 4th ed. Oxford: Oxford University Press; 2008.
2. Irrgang JJ, Anderson AF, Boland AL, Harner CD, Kurosaka M, Neyret P, et al. Development and validation of the International Knee Documentation Committee subjective knee form. Am J Sports Med 2001;29:600–13.
3. Hefti F, Muller W, Jakob RP, Staubli HU. Evaluation of knee ligament injuries with the IKDC form. Knee Surg Sports Traumatol Arthrosc 1993;1:226–34.
4. Anderson AF. Rating scales. In: Fu FH, Harner CD, Vince KL, editors. Knee surgery. Baltimore: Williams & Wilkins; 1994. p. 275–96.
5. Anderson AF, Irrgang JJ, Kocher MS, Mann BJ, Harrast JJ. The International Knee Documentation Committee Subjective Knee Evaluation Form: normative data. Am J Sports Med 2006;34:128–35.
6. Padua R, Bondi R, Ceccarelli E, Bondi L, Romanini E, Zanoli G, et al. Italian version of the International Knee Documentation Committee Subjective Knee Form: cross-cultural adaptation and validation. Arthroscopy 2004;20:819–23.
7. Ahmed KM, Said HG, Ramadan EK, Abd el-Radi M, el-Assal MA. Arabic translation and validation of three knee scores, Lysholm Knee Score (LKS), Oxford Knee Score (OKS), and International Knee Documentation Committee Subjective Knee Form (IKDC) [published erratum appears in SICOT J 2019;5:27]. SICOT J 2019;5:6.
8. Metsavaht L, Leporace G, Riberto M, de Mello Sposito MM, Batista LA. Translation and cross-cultural adaptation of the Brazilian version of the International Knee Documentation Committee Subjective Knee Form: validity and reproducibility. Am J Sports Med 2010;38:1894–9.
9. Fu SN, Chan YH. Translation and validation of Chinese version of International Knee Documentation Committee Subjective Knee Form. Disabil Rehabil 2011;33:1186–9.
10. Jia ZY, Zhang C, Zou Y, Huang X, Xu WD. Translation and validation of the Simplified Chinese version of International Knee Documentation Committee Subjective Knee Form. Arch Orthop trauma Surg 2018;138:1433–41.
11. Fu SN, Chan YH. Translation and validation of Chinese version of International Knee Documentation Committee Subjective Knee Form. Disabil Rehabil 2011;33:1186–9.
12. Jacobsen JS, Knudsen P, Fynbo C, Rolving N, Warming S. Reproducibility and responsiveness of a Danish Pedi-IKDC Subjective Knee Form for children with knee disorders. Scand J Med Sci Sports 2016;26:1408–14.
13. Haverkamp D, Sierevelt IN, Breugem SJ, Lohuis K, Blankevoort L, vanDijk CN. Translation and validation of the Dutch version of the International Knee Documentation Committee Subjective Knee Form. Am J Sports Med 2006;34:1680–4.
14. Kümmel D, Preiss S, Harder LP, Leunig M, Impellizzeri FM. Measurement properties of the German version of the IKDC Subjective Knee Sorm (IKDC-SKF). J Patient Rep Outcomes 2018;2:31.
15. Koumantakis GA, Tsoligkas K, Papoutsidakis A, Ververidis A, Drosos GI. Cross-cultural adaptation and validation of the International Knee Documentation Committee Subjective Knee Form in Greek. J Orthop Traumatol 2016;17:123–9.
16. Kim JG, Ha JK, Lee JY, Seo SS, Choi CH, Lee MC. Translation and validation of the Korean version of the International Knee Documentation Committee Subjective Knee Form. Knee Surg Relat Res 2013;23:106.
17. Todor A, Vermesan D, Haragus H, Patrascu JM, Timar B, Cosma DI. Cross-cultural adaptation and validation of the Romanian International Knee Documentation Committee-Subjective Knee Form. PeerJ 2020;8:e8448.
18. Lertwanich P, Praphruetkit T, Keyurapan E, Lamsam C, Kulthanan T. Validity and reliability of Thai version of the International Knee Documentation Committee Subjective Knee Form. J Med Assoc Thai 2008;91:1218–25.
19. Çelik D, Coşkunsu D, KiliÇoğlu Ö, Ergönül Ö, Irrgang JJ. Translation and cross-cultural adaptation of the International Knee Documentation Committee Subjective Knee Form into Turkish. J Orthop Sports Phys Ther 2014;44:899–909.
20. Kocher MS, Smith JT, Iversen MD, Brustowicz K, Ogunwole O, Andersen J, et al. Reliability, validity, and responsiveness of a modified International Knee Documentation Committee Subjective Knee Form (Pedi-IKDC) in children with knee disorders. Am J Sports Med 2011;39:933–9.
21. Nasreddine AY, Connell PL, Kalish LA, Nelson S, Iversen MD, Anderson AF, et al. The Pediatric International Knee Documentation Committee (pedi-ikdc) Subjective Knee Evaluation Form: normative data. Am J Sports Med 2017;45:527–34.
22. Sabatino MJ, Gans CV, Zynda AJ, Chung JS, Miller SM, Wilson PL, et al. An electronic patient-reported outcomes measurement system in paediatric orthopaedics. J Child Orthop 2019;13:431–7.
23. Crawford K, Briggs KK, Rodkey WG, Steadman JR. Reliability, validity, and responsiveness of the IKDC score for meniscus injuries of the knee. Arthroscopy 2007;23:839–44.
24. Grevnerts HT, Terwee CB, Kvist J. The measurement properties of the IKDC-subjective knee form. Knee Surg Sports Traumatol Arthrosc 2015;23:3698–706.
25. Huang CC, Chen WS, Tsai MW, Wang WT. Comparing the Chinese versions of two knee-specific questionnaires (IKDC and KOOS): reliability, validity, and responsiveness. Health Qual Life Outcomes 2017;15:238.
26. Van de Graaf VA, Wolterbeek N, Scholtes VA, Mutsaerts EL, Poolman RW. Reliability and validity of the IKDC, KOOS, and WOMAC for patients with meniscal injuries. Am J Sports Med 2014;42:1408–16.
27. Higgins LD, Taylor MK, Park D, Ghodadra N, Marchant M, Pietrobon R, et al. Reliability and validity of the International Knee Documentation Committee (IKDC) Subjective Knee Form. Joint Bone Spine 2007;74:594–9.
28. Van der Velden CA, van der Steen MC, Leenders J, van Douveren FQ, Janssen RP, Reijman M. Pedi-IKDC or KOOS-child: which questionnaire should be used in children with knee disorders? [Clinical study]. BMC Musculoskelet Disord 2019;20:240.
29. Agel J, LaPrade RF. Assessment of differences between the modified Cincinnati and International Knee Documentation Committee patient outcome scores: a prospective study. Am J Sports Med 2009;37:2151–7.

30. Van Meer BL, Meuffels DE, Vissers MM, Bierma-Zeinstra SM, Verhaar JA, Terwee CB, et al. Knee injury and Osteoarthritis Outcome Score or International Knee Documentation Committee Subjective Knee Form: which questionnaire is most useful to monitor patients with an anterior cruciate ligament rupture in the short term? [Clinical study]. Arthroscopy 2013;29:701–15.
31. Noorduyn JC, van de Graaf VA, Mokkink LB, Willigenburg NW, Poolman RW, ESCAPE Research Group. Responsiveness and minimal important change of the IKDC of middle-aged and older patients with a meniscal tear. Am J Sports Med 2019;47: 364–71.
32. Greco NJ, Anderson AF, Mann BJ, Cole BJ, Farr J, Nissen CW, et al. Responsiveness of the International Knee Documentation Committee Subjective Knee Form in comparison to the Western Ontario and McMaster Universities Osteoarthritis Index, modified Cincinnati Knee Rating System, and Short Form 36 in patients with focal articular cartilage defects. Am J Sports Med 2010;38:891–902.
33. Irrgang JJ, Anderson AF, Boland AL, Harner CD, Neyret P, Richmond JC, et al. Responsiveness of the International Knee Documentation Committee Subjective Knee Form. Am J Sports Med 2006;34:1567–73.
34. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS): development of a self-administered outcome measure. J Orthop Sports Phys Ther 1998;28:88–96.
35. Paradowski PT, Bergman S, Sunden-Lundius A, Lohmander LS, Roos EM. Knee complaints vary with age and gender in the adult population: population-based reference data for the Knee injury and Osteoarthritis Outcome Score (KOOS). BMC Musculoskelet Disord 2006;7:38.
36. Roos EM, Roos HP, Ekdahl C, Lohmander LS. Knee injury and Osteoarthritis Outcome Score (KOOS): validation of a Swedish version. Scand J Med Sci Sports 1998;8:439–48.
37. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS): validation and comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes 2003;1:17.
38. Zhang QH, Du SX, Zheng GZ, Chang B, Xie D, Lin FX, et al. Reliability, validity, and responsiveness of the Chinese version of the Knee Injury and Osteoarthritis Outcome Score (KOOS) in patients with anterior cruciate ligament reconstruction in mainland China. Z Orthop Unfall 2019;157:42–7.
39. Cheung RT, Ngai SP, Ho KK. Chinese adaptation and validation of the Knee Injury and Osteoarthritis Outcome Score (KOOS) in patients with knee osteoarthritis. Rheumatol Int 2016;36:1449–54.
40. Xie F, Li SC, Roos EM, Fong KY, Lo NN, Yeo SJ, et al. Cross-cultural adaptation and validation of Singapore English and Chinese versions of the Knee injury and Osteoarthritis Outcome Score (KOOS) in Asians with knee osteoarthritis in Singapore. Osteoarthritis Cartilage 2006;14:1098–103.
41. Moutzouri M, Tsoumpos P, Billis E, Papoutsidakis A, Gliatis J. Cross-cultural translation and validation of the Greek version of the Knee Injury and Osteoarthritis Outcome Score (KOOS) in patients with total knee replacement. Disabil Rehabil 2015;37:1477–83.
42. Briem K. Reliability, validity and responsiveness of the Icelandic version of the Knee Injury and Osteoarthritis Outcome Score (KOOS). Laeknabladid 2012;98:403–7.
43. Vaquero J, Longo UG, Forriol F, Martinelli N, Vethencourt R, Denaro V. Reliability, validity and responsiveness of the Spanish version of the Knee Injury and Osteoarthritis Outcome Score (KOOS) in patients with chondral lesion of the knee. Knee Surg Sports Traumatol Arthrosc 2014;22:104–8.
44. Lizaur-Utilla A, Miralles-Muñoz FA, Gonzalez-Parreño S, Lopez-Prats FA. Validation of the Spanish version of the Knee Injury and Osteoarthritis Outcome Score (KOOS) for elderly patients with total knee replacement. J Orthop Res 2019;37:2157–62.
45. De Groot IB, Favejee MM, Reijman M, Verhaar JA, Terwee CB. The Dutch version of the Knee Injury and Osteoarthritis Outcome Score: a validation study. Health Qual Life Outcomes 2008;6:16.
46. 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 Knee Injury and Osteoarthritis Outcome Score (KOOS) in knee osteoarthritis patients. Osteoarthritis Cartilage 2008;16:423–8.
47. Alfadhel SA, Vennu V, Alnahdi AH, Omar MT, Alasmari SH, AlJafri Z, et al. Cross-cultural adaptation and validation of the Saudi Arabic version of the Knee Injury and Osteoarthritis Outcome Score (KOOS). Rheumatol Int 2018;38:1547–55.
48. Cheng AS, Chan KC, Chan SY, Fan MK, Fung MK, Lee OY, et al. Cross-cultural adaptation and validation of the Hong Kong version of the Knee Injury and Osteoarthritis Outcome Score (HK-KOOS) for patients with knee osteoarthritis. Occup Ther Int 2019;2019:8270637.
49. Lyman S, Omori G, Nakamura N, Takahashi T, Tohyama H, Fukui N, et al. Development and validation of a culturally relevant Japanese KOOS. J Orthop Sci 2019;24:514–20.
50. Salavati M, Mazaheri M, Negahban H, Sohani SM, Ebrahimian MR, Ebrahimi I, et al. Validation of a Persian-version of Knee Injury and Osteoarthritis Outcome Score (KOOS) in Iranians with knee injuries. Osteoarthritis Cartilage 2008;16:1178–82.
51. Goncalves RS, Cabri J, Pinheiro JP, Ferreira PL. Cross-cultural adaptation and validation of the Portuguese version of the Knee Injury and Osteoarthritis Outcome Score (KOOS). Osteoarthritis Cartilage 2009;17:1156–62.
52. Chaipinyo K. Test-retest reliability and construct validity of Thai version of Knee Osteoarthritis Outcome Score (KOOS). Thai J Phys Ther 2009;31:67–76.
53. Paker N, Bugdayci D, Sabirli F, Ozel S, Ersoy S. Knee Injury and Osteoarthritis Outcome Score: reliability and validation of the Turkish version. Turkiye Klinikleri J Med Sci 2007;27:350–6.
54. Bekkers JE, de Windt TS, Raijmakers NJ, Dhert WJ, Saris DB. Validation of the Knee Injury and Osteoarthritis Outcome Score (KOOS) for the treatment of focal cartilage lesions. Osteoarthritis Cartilage 2009;17:1434–9.
55. Steinhoff AK, Bugbee WD. Knee Injury and Osteoarthritis Outcome Score has higher responsiveness and lower ceiling effect than Knee Society Function Score after total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc 2016;24:2627–33.
56. Örtqvist M, Roos EM, Broström EW, Janarv PM, Iversen MD. Development of the knee Injury and osteoarthritis outcome score for children (KOOS-Child) comprehensibility and content validity. Acta Orthop 2012;83:666–73.
57. Ortqvist M, Iversen MD, Janarv PM, Broström EW, Roos EM. Psychometric properties of the Knee Injury and Osteoarthritis Outcome Score for Children (KOOS-Child) in children with knee disorders. Br J Sports Med 2014;48:1437–46.
58. Collins NJ, Prinsen CA, Christensen R, Bartels EM, Terwee CB, Roos EM. Knee Injury and Osteoarthritis Outcome Score (KOOS): systematic review and meta-analysis of measurement properties. Osteoarthritis Cartilage 2016;24:1317–29.
59. Salavati M, Akhbari B, Mohammadi F, Mazaheri M, Khorrami M. Knee Injury and Osteoarthritis Outcome Score (KOOS): reliability and validity in competitive athletes after anterior cruciate ligament reconstruction. Osteoarthritis Cartilage 2011;19:406–10.
60. Comins J, Brodersen J, Krogsgaard M, Beyer N. Rasch analysis of the Knee Injury and Osteoarthritis Outcome Score (KOOS): a statistical re-evaluation. Scand J Med Sci Sports 2008;18:336–5.

61. Peer MA, Lane J. The Knee Injury and Osteoarthritis Outcome Score (KOOS): a review of its psychometric properties in people undergoing total knee arthroplasty. J Orthop Sports Phys Ther 2013;43:20–8.
62. Singh JA1, Luo R, Landon GC, Suarez-Almazor M. Reliability and clinically important improvement thresholds for osteoarthritis pain and function scales: a multicenter study. J Rheumatol 2014;41:509–15.
63. Perruccio AV, Stefan Lohmander L, Canizares M, Tennant A, Hawker GA, Conaghan PG, et al. The development of a short measure of physical function for knee OA KOOS-Physical Function Shortform (KOOS-PS): an OARSI/OMERACT initiative. Osteoarthritis Cartilage 2008;16:542–50.
64. Ornetti P, Perruccio AV, Roos EM, Lohmander LS, Davis AM, Maillefert JF. Psychometric properties of the French translation of the reduced KOOS and HOOS (KOOS-PS and HOOS-PS). Osteoarthritis Cartilage 2009;17:1604–8.
65. Goncalves RS, Cabri J, Pinheiro JP, Ferreira PL, Gil J. Reliability, validity and responsiveness of the Portuguese version of the Knee Injury and Osteoarthritis Outcome Score-Physical Function Shortform (KOOS-PS). Osteoarthritis Cartilage 2010;18:372–6.
66. Gul ED, Yilmaz O, Bodur H. Reliability and validity of the Turkish version of the Knee Injury and Osteoarthritis Outcome Score-Physical Function short-form (KOOS-PS). J Back Musculoskelet Rehabil 2013;26:461–6.
67. Davis AM, Perruccio AV, Canizares M, Hawker GA, Roos EM, Maillefert JF, et al. Comparative, validity and responsiveness of the HOOS-PS and KOOS-PS to the WOMAC physical function subscale in total joint replacement for osteoarthritis. Osteoarthritis Cartilage 2009;17:843–7.
68. Franchignoni F, Salaffi F, Giordano A, Carotti M, Ciapetti A, Ottonello M. Rasch analysis of the 22 knee injury and osteoarthritis outcome score–physical function items in Italian patients with knee osteoarthritis. Arch Phys Med Rehabil 2013;94:480–7.
69. Mahler E, Cuperus N, Bijlsma J, Vliet Vlieland T, van den Hoogen F, den Broeder AA, et al. Responsiveness of four patient-reported outcome measures to assess physical function in patients with knee osteoarthritis. Scand J Rheumatol 2016;45:518–27.
70. Harris KK, Dawson J, Jones LD, Beard DJ, Price AJ. Extending the use of PROMs in the NHS—using the Oxford Knee Score in patients undergoing non-operative management for knee osteoarthritis: a validation study. BMJ Open 2013;3:e003365.
71. Irrgang JJ, Snyder-Mackler L, Wainner RS, Fu FH, Harner CD. Development of a patient-reported measure of function of the knee. J Bone Joint Surg Am 1998;80:1132–45.
72. Marx RG. Knee rating scales. Arthroscopy 2003;19:1103–8.
73. Piva SR, Gil AB, Moore CG, Fitzgerald GK. Responsiveness of the activities of daily living scale of the knee outcome survey and numeric pain rating scale in patients with patellofemoral pain. J Rehabil Med 2009;41:129–35.
74. Impellizzeri FM, Mannion AF, Leunig M, Bizzini M, Naal FD. Comparison of the reliability, responsiveness, and construct validity of 4 different questionnaires for evaluating outcomes after total knee arthroplasty. J Arthroplasty 2011;26:861–9.
75. Marx RG, Jones EC, Allen AA, Altchek DW, O'Brien SJ, Rodeo SA, et al. Reliability, validity, and responsiveness of four knee outcome scales for athletic patients. J Bone Joint Surg Am 2001;83:1459–69.
76. Irrgang JJ. Development of a health related quality of life instrument to assess physical function related to pathology and impairment of the knee. Pittsburgh: University of Pittsburgh; 1999.
77. Algarni AD, Alrabai HM, Al-Ahaideb A, Kachanathu SJ, AlShammari SA. Arabic translation, cultural adaptation, and validation study of Knee Outcome Survey: Activities of Daily Living Scale (KOS-ADLS). Rheumatol Int 2017;37:1585–9.
78. Bouzubar FF, Aljadi SH, Alotaibi NM, Irrgang JJ. Cross-cultural adaptation and validation of the Arabic version of the Knee Outcome Survey-Activities for Daily Living scale. Disabil Rehabil 2018;40:1817–28.
79. Jia ZY, Wang W, Nian XW, Zhang XX, Huang ZP, Cui J, et al. Cross-cultural adaptation and validation of the simplified Chinese version of the Knee Outcome Survey Activities of Daily Living Scale. Arthroscopy 2016;32:2009–16.
80. Roy JS, Esculier JF, Maltais DB. Translation, cross-cultural adaptation and validation of the French version of the Knee Outcome Survey-Activities of Daily Living Scale. Clin Rehabil 2014;28:614–23.
81. Bizzini M, Gorelick M. Development of a German version of the Knee Outcome Survey for Daily Activities. Arch Orthop Trauma Surg 2007;127:781–9.
82. Goncalves R, Cabri J, Pinheiro J. Cross-cultural adaptation and validation of the Portuguese version of the Knee Outcome Survey Activities of Daily Living Scale (KOS-ADLS). Clin Rheumatol 2008;27:1445–9.
83. Szczepanik M, Bejer A, Snela S, Szymczyk D, Jabłoński J, Majewska J. Polish cross-cultural adaptation and validation of the Knee Outcome Survey Activities of Daily Living Scale (KOS-ADLS) in patients undergoing total knee arthroplasty. Med Sci Monit 2018;24:5309–19.
84. Evcik D, Ay S, Ege A, Turel A, Kavuncu V. Adaptation and validation of Turkish version of the Knee Outcome Survey-Activities for Daily Living Scale. Clin Orthop Relat Res 2009;467:2077–82.
85. Kapreli E, Panelli G, Strimpakos N, Billis E, Zacharopoulos A, Athanasopoulos S. Cross-cultural adaptation of the Greek version of the Knee Outcome Survey-Activities of Daily Living Scale (KOS-ADLS). Knee 2011;18:424–7.
86. Esculier JF, Roy JS, Bouyer LJ. Psychometric evidence of self-reported questionnaires for patellofemoral pain syndrome: a systematic review. Disabil Rehabil 2013;35:2181–90.
87. Williams VJ, Piva SR, Irrgang JJ, Crossley C, Fitzgerald GK. Comparison of reliability and responsiveness of patient-reported clinical outcome measures in knee osteoarthritis rehabilitation. J Orthop Sports Phys Ther 2012;42:716–23.
88. Lysholm J, Gillquist J. Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. Am J Sports Med 1982;10:150–4.
89. Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. Clin Orthop Relat Res 1985;198:43–9.
90. Briggs KK, Lysholm J, Tegner Y, Rodkey WG, Kocher MS, Steadman JR. The reliability, validity, and responsiveness of the Lysholm Score and Tegner Activity Scale for anterior cruciate ligament injuries of the knee: 25 years later. Am J Sports Med 2009;37:890–7.
91. Hoher J, Bach T, Munster A, Bouillon B, Tiling T. Does the mode of data collection change results in a subjective knee score? Self-administration versus interview. Am J Sports Med 1997;25:642–7.
92. Smith HJ, Richardson JB, Tennant A. Modification and validation of the Lysholm Knee Scale to assess articular cartilage damage. Osteoarthritis Cartilage 2009;17:53–8.
93. Nguyen J, Marx R, Hidaka C, Wilson S, Lyman S. Validation of electronic administration of knee surveys among ACL-injured patients. Knee Surg Sports Traumatol Arthrosc 2017;25:3116–22.
94. Kose O, Deniz G, Ozcan H, Guler F. A comparison of telephone interview versus on-site completion of Lysholm Knee Score in patients who underwent arthroscopic ACL reconstruction: are the results equivalent? [Clinical review study]. Eur J Orthop Surg Traumatol 2015;25:1069–72.
95. Briggs KK, Steadman JR, Hay CJ, Hines SL. Lysholm score and Tegner activity level in individuals with normal knees. Am J Sports Med 2009;37:898–901.

96. Demirdjian AM, Petrie SG, Guanche CA, Thomas KA. The outcomes of two knee scoring questionnaires in a normal population. Am J Sports Med 1998;26:46–51.
97. Wang W, Liu L, Chang X, Jia ZY, Zhao JZ, Xu WD. Cross-cultural translation of the Lysholm Knee Score in Chinese and its validation in patients with anterior cruciate ligament injury. BMC Musculoskelet Disord 2016;17:436.
98. Eshuis R, Lentjes GW, Tegner Y, Wolterbeek N, Veen MR. Dutch translation and cross-cultural adaptation of the Lysholm Score and Tegner Activity Scale for patients with anterior cruciate ligament injuries. J Orthop Sports Phys Ther 2016;46:976–83.
99. Cerciello S, Corona K, Morris BJ, Visonà E, Maccauro G, Maffulli N, et al. Cross-cultural adaptation and validation of the Italian versions of the Kujala, Larsen, Lysholm and Fulkerson Score in patients with patellofemoral disorders. J Orthop Traumatol 2018;19:18.
100. Swanenburg J, Koch PP, Meier N, Wirth B. Function and activity in patients with knee arthroplasty: validity and reliability of a German version of the Lysholm Score and the Tegner Activity Scale. Swiss Med Wkly 2014;144:w13976.
101. Arroyo-Morales M, Martin-Alguacil J, Lozano-Lozano M, Cuesta-Vargas AI, Fernández-Fernández AJ, González JA, et al. The Lysholm Score: cross cultural validation and evaluation of psychometric properties of the Spanish version. PLoS One 2019;14:e0221376.
102. Oretorp N, Gillquist J, Liljedahl SO. Long term results of surgery for non-acute anteromedial rotatory instability of the knee. Acta Orthop Scand 1979;50:329–36.
103. Briggs KK, Kocher MS, Rodkey WG, Steadman JR. Reliability, validity and responsiveness of the Lysholm Knee Score and the Tegner Activity Scale for patients with meniscal injury of the knee. J Bone Joint Surg Am 2006;88:698–705.
104. Heintjes EM, Bierma-Zeinstra SM, Berger MY, Koes BW. Lysholm Scale and WOMAC Index were responsive in prospective cohort of young general practice patients. J Clin Epidemiol 2008;61:481–8.
105. Kocher MS, Steadman JR, Briggs KK, Sterett WI, Hawkins RJ. Reliability, validity, and responsiveness of the Lysholm Knee Scale for various chondral disorders of the knee. J Bone Joint Surg Am 2004;86:1139–45.
106. Paxton EW, Fithian DC, Stone ML, Silva P. The reliability and validity of knee-specific and general health instruments in assessing acute patellar dislocation outcomes. Am J Sports Med 2003;31:487–92.
107. Sgaglione NA, del Pizzo W, Fox JM, Friedman MJ. Critical analysis of knee ligament rating systems. Am J Sports Med 1995;23:660–7.
108. Howard JS, Lattermann C, Hoch JM, Mattacola CG, Medina McKeon JM. Comparing responsiveness of six common patient-reported outcomes to changes following autologous chondrocyte implantation: a systematic review and meta-analysis of prospective studies. Cartilage 2013;4:97–110.
109. Hambly K. The use of the Tegner Activity Scale for articular cartilage repair of the knee: a systematic review. Knee Surg Sports Traumatol Arthrosc 2011;19:604–14.
110. Frobell RB, Roos EM, Roos HP, Ranstam J, Lohmander LS. A randomized trial of treatment for acute anterior cruciate ligament tears. N Engl J Med 2010;363:331–42.
111. Naal FD, Impellizzeri FM, Leunig M. Which is the best activity rating scale for patients undergoing total joint arthroplasty? [Clinical study]. Clin Orthop Relat Res 2009;467:958–65.
112. Flosadottir V, Roos EM, Ageberg E. Translation, cross-cultural adaptation, and validation of the Activity Rating Scale for disorders of the knee. Orthop J Sports Med 2017;5:2325967117729361.
113. Mostafaee N, Negahban H, Shaterzadeh Yazdi MJ, Goharpey S, Mehravar M, Pirayeh N. Responsiveness of a Persian version of Knee Injury and Osteoarthritis Outcome Score and Tegner Activity Scale in athletes with anterior cruciate ligament reconstruction following physiotherapy treatment. Physiother Theory Pract 2018;1–8.
114. Huang H, Zhang D, Jiang Y, Yang J, Feng T, Gong X, et al. Translation, validation and cross-sultural adaptation of a simplified-Chinese version of the Tegner Activity Score in Chinese patients with anterior cruciate ligament injury. PLoS One 2016;11:e0155463.
115. Sabatino MJ, Gans CV, Zynda AJ, Chung JS, Miller SM, Wilson PL, et al. An electronic patient-reported outcomes measurement system in paediatric orthopaedics. J Child Orthop 2019;13:431–7.
116. Letchford R, Sparkes V, van Deursen RW. Assessing participation in the ACL injured population: selecting a patient reported outcome measure on the basis of measurement properties. Knee 2015;22:262–9.
117. Frobell RB, Svensson E, Gothrick M, Roos EM. Self-reported activity level and knee function in amateur football players: the influence of age, gender, history of knee injury and level of competition. Knee Surg Sports Traumatol Arthrosc 2008;16:713–9.
118. 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.
119. 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.
120. Harris KK, Dawson J, Jones LD, Beard DJ, Price AJ. Extending the use of PROMs in the NHS–using the Oxford Knee Score in patients undergoing non-operative management for knee osteoarthritis: a validation study. BMJ Open 2013;3:e003365.
121. Alghadir A, Anwer S, Iqbal ZA, Alsanawi HA. Cross-cultural adaptation, reliability and validity of the Arabic version of the reduced Western Ontario and McMaster Universities Osteoarthritis index in patients with knee osteoarthritis. Disabil Rehabil 2016;38:689–94.
122. Xie F, Li SC, Lo NN, Yeo SJ, Yang KY, Yeo W, et al. Cross-cultural adaptation and validation of Singapore English and Chinese Versions of the Oxford Knee Score (OKS) in knee osteoarthritis patients undergoing total knee replacement. Osteoarthritis Cartilage 2007;15:1019–24.
123. Cheung RT, Ngai SP, Ho KK. Chinese translation and validation of the Oxford Knee Scale for patients with knee osteoarthritis. Hong Kong Physiother J 2017;37:46–9.
124. Lin K, Bao L, Wang J, Fujita K, Makimoto K, Liao X. Validation of the Chinese (Mandarin) version of the Oxford Knee Score in patients with knee osteoarthritis. Clin Orthop Relat Res 2017;475:2992–3004.
125. Reito A, Järvistö A, Jämsen E, Skyttä E, Remes V, Huhtala H, et al. Translation and validation of the 12-item Oxford Knee Score for use in Finland. BMC Musculoskelet Disord 2017;18:74.
126. Naal F, Impellizzeri F, Sieverding M, Loibl M, von Knoch F, Mannion A, et al. The 12-item Oxford Knee Score: cross-cultural adaptation into German and assessment of its psychometric properties in patients with osteoarthritis of the knee. Osteoarthritis Cartilage 2009;17:49–52.
127. Takeuchi R, Sawaguchi T, Nakamura N, Ishikawa H, Saito T, Goldhahn S. Cross-cultural adaptation and validation of the Oxford 12-item Knee Score in Japanese. Arch Orthop Trauma Surg 2011;131:247–54.
128. Eun IS, Kim OG, Kim CK, Lee HS, Lee JS. Validation of the Korean version of the Oxford Knee Score in patients undergoing total knee arthroplasty. Clin Orthop Relat Res 2013;471:600–5.
129. Ebrahimzadeh MH, Makhmalbaf H, Birjandinejad A, Keshtan FG, Hoseini HA, Mazloumi SM. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) in Persian speaking patients with knee osteoarthritis. Arch Bone Jt Surg 2014;2:57–62.

130. Ebrahimzadeh MH, Makhmalbaf H, Birjandinejad A, Soltani-Moghaddas SH. Cross-cultural adaptation and validation of the Persian version of the Oxford Knee Score in patients with knee osteoarthritis. Iran J Med Sci 2014;39:529–35.
131. Gonçalves RS, Tomás AM, Martins DI. Cross-cultural adaptation and validation of the Portuguese version of the Oxford Knee Score (OKS). Knee 2012;19:344–347.
132. Martín-Fernández J, García-Maroto R, Sánchez-Jiménez FJ, Bau-González A, Valencia-García H, Gutiérrez-Teira B, et al. Validation of the Spanish version of the Oxford Knee Score and assessment of its utility to characterize quality of life of patients suffering from knee osteoarthritis: a multicentric study. Health Qual Life Outcomes 2017;15:186.
133. Dunbar M, Robertsson O, Ryd L, Lidgren L. Translation and validation of the Oxford-12 item Knee Score for use in Sweden. Acta Orthop Scand 2000;71:268–74.
134. Charoencholvanich K, Pongcharoen B. Oxford Knee Score and SF-36: translation & reliability for use with total knee arthroscopy patients in Thailand. J Med Assoc Thai 2005;88:1194–202.
135. Tuğay BU, Tuğay N, Güney H, Kınıklı Gİ, Yüksel İ, Atilla B. Oxford Knee Score: cross-cultural adaptation and validation of the Turkish version in patients with osteoarthritis of the knee. Acta Orthop Traumatol Turc 2016;50:198–206.
136. Jenny JY, Diesinger Y. The Oxford Knee Score: compared performance before and after knee replacement. Orthop Traumatol Surg Res 2012;98:409–12.
137. Ko Y, Lo NN, Yeo SJ, Yang KY, Yeo W, Chong HC, et al. Comparison of the responsiveness of the SF-36, the Oxford Knee Score, and the Knee Society Clinical Rating System in patients undergoing total knee replacement. Qual Life Res 2013;22:2455–9.
138. Harris K, Dawson J, Doll H, Field RE, Murray DW, Fitzpatrick R, et al. Can pain and function be distinguished in the Oxford Knee Score in a meaningful way? An exploratory and confirmatory factor analysis. Qual Life Res 2013;22:2561–8.
139. Bream E, Charman SC, Clift B, Murray D, Black N. Relationship between patients' and clinicians' assessments of health status before and after knee arthroplasty. Qual Saf Health Care 2010;19:e6.
140. Conaghan PG, Emerton M, Tennant A. Internal construct validity of the Oxford Knee Scale: evidence from Rasch measurement. Arthritis Rheum 2007;57:1363–7.
141. Ko Y, Lo NN, Yeo SJ, Yang KY, Yeo W, Chong HC, et al. Rasch analysis of the Oxford Knee Score. Osteoarthritis Cartilage 2009;17:1163–9.
142. Garratt AM, Brealey S, Gillespie WJ, DAMASK Trial Team. Patient-assessed health instruments for the knee: a structured review. Rheumatology (Oxford) 2004;43:1414–23.
143. Goldhahn S, Takeuchi R, Nakamura N, Nakamura R, Sawaguchi T. Responsiveness of the Knee Injury and Osteoarthritis Outcome Score (KOOS) and the Oxford Knee Score (OKS) in Japanese patients with high tibial osteotomy. J Orthop Sci 2017;22:862–7.
144. Moonot P, Medalla GA, Matthews D, Kalairajah Y, Field RE. Correlation between the Oxford Knee and American Knee Society scores at mid-term follow-up. J Knee Surg 2009;22:226–30.
145. Maempel JF, Clement ND, Brenkel IJ, Walmsley PJ. Range of movement correlates with the Oxford Knee Score after total knee replacement: a prediction model and validation. Knee 2016;23:511–6.
146. 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.
147. 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.
148. 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.
149. Medalla GA, Moonot P, Peel T, Kalairajah Y, Field RE. Cost-benefit comparison of the Oxford Knee Score and the American Knee Society Score in measuring outcome of total knee arthroplasty. J Arthroplasty 2009;24:652–6.
150. Maempel JF, Clement ND, Brenkel IJ, Walmsley PJ. Validation of a prediction model that allows direct comparison of the Oxford Knee Score and American Knee Society clinical rating system. Bone Joint J 2015;97-B:503–9.
151. Bellamy N. WOMAC Osteoarthritis Index user guide. London (Ontario, Canada): University of Western Ontario; 1995.
152. Bellamy N. WOMAC Osteoarthritis Index user guide. Version V. Brisbane (Australia): CONROD, The University of Queensland; 2002.
153. WOMAC-AUSCAN-osteoarthritis global index. 2016. URL: http://www.womac.com.
154. 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.
155. Baron G, Tubach F, Ravaud P, Logeart I, Dougados M. Validation of a short form of the Western Ontario and McMaster Universities Osteoarthritis Index function subscale in hip and knee osteoarthritis. Arthritis Rheum 2007;57:633–8.
156. Whitehouse SL, Lingard EA, Katz JN, Learmonth ID. Development and testing of a reduced WOMAC function scale. J Bone Joint Surg Br 2003;85:706–11.
157. Yang KG, Raijmakers NJ, Verbout AJ, Dhert WJ, Saris DB. Validation of the short-form WOMAC function scale for the evaluation of osteoarthritis of the knee. J Bone Joint Surg Br 2007;89:50–6.
158. Shen ZD, Yu HM, Wang JT, Shi GY, Sun Y. Modified Western Ontario and McMaster University Osteoarthritis Index Scale used in patients with knee osteoarthritis. Zhonghua Yi Xue Za Zhi 2019;99:537–41.
159. 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.
160. Bellamy N, Campbell J, Hill J, Band P. A comparative study of telephone versus onsite completion of the WOMAC 3.0 Osteoarthritis Index. J Rheumatol 2002;29:783–6.
161. 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.
162. Theiler R, Speilberger J, Bischoff H, 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.
163. Marsh JD, Bryant DM, Macdonald SJ, Naudie DD. Patients respond similarly to paper and electronic versions of the WOMAC and SF-12 following total joint arthroplasty. J Arthroplasty 2014;29:670–3.
164. 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.
165. 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.

166. Rabbani MG, Haq SA, Bellamy N, Islam MN, Choudhury MR, Naheed A, et al. Development, linguistic and clinimetric validation of the WOMAC VA3.01 Bangla for Bangladesh Index. Rheumatol Int 2015;35:997–1003.
167. 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.
168. Symonds T, Hughes B, Liao S, Ang Q, Bellamy N. Validation of the Chinese Western Ontario and McMaster Universities Osteoarthritis Index in patients from mainland China with osteoarthritis of the knee. Arthritis Care Res (Hoboken) 2015;67:1553–60.
169. 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.
170. 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.
171. Papathanasiou G, Stasi S, Oikonomou L, Roussou I, Papageorgiou E, Chronopoulos E, et al. Clinimetric properties of WOMAC Index in Greek knee osteoarthritis patients: comparisons with both self-reported and physical performance measures. Rheumatol Int 2015;35:115–23.
172. 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.
173. 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.
174. 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.
175. 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.
176. 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.
177. Nakarmi S, Haq SA, Vaidya B. Translation, validation and cross-cultural adaptation of the Nepali version of WOMAC® LK 3.1. Int J Rheum Dis 2019;22:1877–83.
178. Lage PT, Machado LA, Barreto SM, de Figueiredo RC, Telles RW. Measurement properties of Portuguese-Brazil Western Ontario and McMaster Universities osteoarthritis index (WOMAC) for the assessment of knee complaints in Brazilian adults. Rheumatol Int 2020;40:233–42.
179. 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.
180. Escobar A, Quintana JM, Bilbao A, Azkarate J, Guenaga 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.
181. Roos EM, Klassbo M, Lohmander LS. Western Ontario and McMaster Universities. WOMAC osteoarthritis index: reliability, validity, and responsiveness in patients with arthroscopically assessed osteoarthritis. Scand J Rheumatol 1999;28:210–5.
182. Soderman 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.
183. 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.
184. 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.
185. Tuzun EH, Eker L, Aytar A, Daskapan A, Bayramoglu M. Acceptability, reliability, validity and responsiveness of the Turkish version of WOMAC osteoarthritis index. Osteoarthritis Cartilage 2005;13:28–33.
186. Escobar A, Quintana JM, Bilbao A, Arostegui I, Lafuente I, Vidaurretal I. Responsiveness and clinically important differences for the WOMAC and SF-36 after total knee replacement. Osteoarthritis Cartilage 2007;15:273–80.
187. Giesinger K, Hamilton DF, Jost B, Holzner B, Giesinger JM. Comparative responsiveness of outcome measures for total knee arthroplasty. Osteoarthritis Cartilage 2014;22:184–9.
188. Ayala A, Bilbao A, Garcia-Perez S, Escobar A, Forjaz MJ. Scale invariance and longitudinal stability of the Physical Functioning Western Ontario and MacMaster Universities Osteoarthritis Index using the Rasch model. Rheumatol Int 2018;38:473–9.
189. White DK, Master H. Patient-reported measures of physical function in knee osteoarthritis. Rheum Dis Clin North Am 2016;42:239–52.
190. Riddle DL, Stratford PW. Unilateral vs bilateral symptomatic knee osteoarthritis: associations between pain intensity and function. Rheumatology (Oxford) 2013;52:2229–37.
191. Brazier JE, Harper R, Munro J, Walters SJ, Snaith ML. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. Rheumatology (Oxford) 1999;38:870–7.
192. 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.
193. Gill SD, de Morton NA, Mc Burney H. An investigation of the validity of six measures of physical function in people awaiting joint replacement surgery of the hip or knee. Clin Rehabil 2012;26:945–51.
194. Davis AM, Badley EM, Beaton DE, Kopec J, Wright JG, Young NL, et al. Rasch analysis of the Western Ontario McMaster (WOMAC) Osteoarthritis Index: results from community and arthroplasty samples. J Clin Epidemiol 2003;56:1076–83.
195. Wolfe F, Kong SX. Rasch analysis of the Western Ontario McMaster questionnaire (WOMAC) in 2205 patients with osteoarthritis, rheumatoid arthritis, and fibromyalgia. Ann Rheum Dis 1999;58:563–8.
196. Giesinger JM, Kuster MS, Behrend H, Giesinger K. Association of psychological status and patient-reported physical outcome measures in joint arthroplasty: a lack of divergent validity. Health Qual Life Outcomes 2013;11:64.
197. Tevald MA, Murray A, Luc BA, Lai K, Sohn D, Pietrosimone B. Hip abductor strength in people with knee osteoarthritis: a cross-sectional study of reliability and association with function. Knee 2016;23:57–62.
198. Trudeau J, van Inwegen R, Eaton T, Bhat G, Paillard F, Ng D, et al. Assessment of pain and activity using an electronic pain diary and actigraphy device in a randomized, placebo-controlled crossover trial of celecoxib in osteoarthritis of the knee. Pain Pract 2015;15:247–55.

199. 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.
200. French HP, Fitzpatrick M, FitzGerald O. Responsiveness of physical function outcomes following physiotherapy intervention for osteoarthritis of the knee: an outcome comparison study. Physiotherapy 2011;97:302–8.
201. Bolink SA, Grimm B, Heyligers IC. Patient-reported outcome measures versus inertial performance-based outcome measures: a prospective study in patients undergoing primary total knee arthroplasty. Knee 2015;22:618–23.
202. Parkes MJ, Callaghan MJ, Tive L, Lunt M, Felson DT. Responsiveness of single versus composite measures of pain in knee osteoarthritis. J Rheumatol 2018;45:1308–15.
203. Jogi P, Spaulding SJ, Zecevic AA, Overend TJ, Kramer JF. Comparison of the original and reduced versions of the Berg Balance Scale and the Western Ontario and McMaster Universities Osteoarthritis Index in patients following hip or knee arthroplasty. Physiother Can 2011;63:107–14.
204. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. Ann Rheum Dis 2005;64:34–7.
205. Abbott JH, Hobbs C, Gwynne-Jones D, MOA Trial Team. The ShortMAC: minimum important change of a reduced version of the Western Ontario and McMaster Universities Osteoarthritis Index. J Orthop Sports Phys Ther 2018;48:81–6.
206. Baranova IV. The use of the functional state of the joints for the estimation of the effectiveness of the application of oxygen/ozone therapy for the rehabilitative treatment of the patients suffering from knee arthritis. Vopr Kurortol Fizioter Lech Fiz Kult 2018;95:42–8.
207. Kemp MA, Metcalfe AJ, Sayers A, Wylde V, Eldridge JD, Blom AW. Does overstuffing of the patellofemoral joint in total knee arthroplasty have a significant effect on postoperative outcomes? [Prospective cohort study]. Knee 2018;25:874–81.
208. Tive L, Bello AE, Radin D, Schnitzer TJ, Nguyen H, Brown MT, et al. Pooled analysis of tanezumab efficacy and safety with subgroup analyses of phase III clinical trials in patients with osteoarthritis pain of the knee or hip. J Pain Res 2019;12:975–95.
209. Quicke JG, Foster NE, Croft PR, Ogollah RO, Holden MA. Change in physical activity level and clinical outcomes in older adults with knee pain: a secondary analysis from a randomised controlled trial. BMC Musculoskelet Disord 2018;19:59.
210. Lopez HL, Habowski SM, Sandrock JE, Raub B, Kedia A, Bruno EJ, et al. Effects of dietary supplementation with a standardized aqueous extract of Terminalia chebula fruit (AyuFlex®) on joint mobility, comfort, and functional capacity in healthy overweight subjects: a randomized placebo-controlled clinical trial. BMC Complement Altern Med 2017;17:475.
211. Wirth W, Hunter DJ, Nevitt MC, Sharma L, Kwoh CK, Ladel C, et al. Predictive and concurrent validity of cartilage thickness change as a marker of knee osteoarthritis progression: data from the Osteoarthritis Initiative. Osteoarthritis Cartilage 2017;25:2063–71.
212. Marx RG, Stump TJ, Jones EC, Wickiewicz TL, Warren RF. Development and evaluation of an activity rating scale for disorders of the knee. Am J Sports Med 2001;29:213–8.
213. Nguyen J, Marx R, Hidaka C, Wilson S, Lyman S. Validation of electronic administration of knee surveys among ACL-injured patients. Knee Surg Sports Traumatol Arthrosc 2017;25:3116–22.
214. Iversen MD, von Heideken J, Farmer E, Rihm J, Heyworth BE, Kocher MS. Validity and comprehensibility of physical activity scales for children with sport injuries. J Pediatr Orthop 2016;36:278–83.
215. Mostafaee N, Negahban H, Shaterzadeh Yazdi MJ, Goharpey S, Mehravar M, Pirayeh N. Responsiveness of a Persian version of Knee Injury and Osteoarthritis Outcome Score and Tegner Activity Scale in athletes with anterior cruciate ligament reconstruction following physiotherapy treatment. Physiother Theory Pract 2018;1–8.
216. Shirazi CP, Israel HA, Kaar SG. Is the Marx Activity Scale reliable in patients younger than 18 years? [Prospective cohort study]. Sports Health 2016;8:145–8.
217. Cameron KL, Peck KY, Thompson BS, Svoboda SJ, Owens BD, Marshall SW. Reference values for the Marx Activity Rating Scale in a young athletic population: history of knee ligament injury is associated with higher scores. Sports Health 2015;7:403–8.
218. Wright RW. Knee injury outcomes measures. J Am Acad Orthop Surg 2009;17:31–9.
219. 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.
220. HealthMeasures. PROMIS instrument development and validation scientific standards. 2013. URL: http://www.healthmeasures.net/images/PROMIS/PROMISStandards_Vers2.0_Final.pdf.
221. Gulbrandsen TR, Khazi ZM, Bollier M, Wolf B, Larson C, Duchman K, et al. Preoperative performance of Patient-Reported Outcomes Measurement Information System in patients with meniscal root tears. J Knee Surg 2019. E-pub ahead of print.
222. Carender CN, Bollier MJ, Wolf BR, Duchman KR, An Q, Westermann RW. Preoperative performance of PROMIS in patients with patellofemoral malalignment and chondral disease. Orthop J Sports Med 2019;7:2325967119855001.
223. Trasolini NA, Korber S, Gipsman A, San AE, Weber AE, Hatch GF. Performance of PROMIS computer adaptive testing as compared with established instruments for multiple-ligament knee injuries. Orthop J Sports Med 2019;7:2325967119867419.
224. Driban JB, Morgan N, Price LL, Cook KF, Wang C. Patient-Reported Outcomes Measurement Information System (PROMIS) instruments among individuals with symptomatic knee osteoarthritis: a cross-sectional study of floor/ceiling effects and construct validity. BMC Musculoskelet Disord 2015;16:253.
225. Scott EJ, Westermann R, Glass NA, Hettrich C, Wolf BR, Bollier MJ. Performance of the PROMIS in patients after anterior cruciate ligament reconstruction. Orthop J Sports Med 2018;6:2325967118774509.
226. Bartlett SJ, Orbai AM, Duncan T, DeLeon E, Ruffing V, Clegg-Smith K, et al. Reliability and validity of selected PROMIS measures in people with rheumatoid arthritis. PloS One 2015;10:e0138543.
227. Hung M, Clegg DO, Greene T, Weir C, Saltzman CL. A lower extremity physical function computerized adaptive testing instrument for orthopaedic patients. Foot Ankle Int 2012;33:326–35.
228. Hung M, Baumhauer JF, Latt LD, Saltzman CL, SooHoo NF, Hunt KJ, et al. Validation of PROMIS® Physical Function computerized adaptive tests for orthopaedic foot and ankle outcome research. Clin Orthop Relat Res 2013;471:3466–74.
229. Oude Voshaar MA, Ten Klooster PM, Glas CA, Vonkeman HE, Taal E, Krishnan E, et al. Validity and measurement precision of the PROMIS physical function item bank and a content validity-driven 20-item short form in rheumatoid arthritis compared with traditional measures. Rheumatology (Oxford) 2015;54:2221–9.
230. Gelhorn HL, Ye X, Speck RM, Tong S, Healey JH, Bukata SV, et al. The measurement of physical functioning among patients with tenosynovial giant cell tumor (TGCT) using the Patient-Reported Outcomes Measurement Information System (PROMIS). J Patient Rep Outcomes 2019;3:6.

231. Crins MH, Terwee CB, Ogreden O, Schuller W, Dekker P, Flens G, et al. Differential item functioning of the PROMIS physical function, pain interference, and pain behavior item banks across patients with different musculoskeletal disorders and persons from the general population. Qual Life Res 2019;28:1231–43.

232. Kenney RJ, Houck J, Giordano BD, Baumhauer JF, Herbert M, Maloney MD. Do Patient Reported Outcome Measurement Information System (PROMIS) scales demonstrate responsiveness as well as disease-specific scales in patients undergoing knee arthroscopy? [Prospective cohort study]. Am J Sports Med 2019;47:1396–403.

233. Fries J, Rose M, Krishnan E. The PROMIS of better outcome assessment: responsiveness, floor and ceiling effects, and Internet administration. J Rheumatol 2011;38:1759–64.

234. Lee AC, Driban JB, Price LL, Harvey WF, Rodday AM, Wang C. Responsiveness and minimally important differences for 4 Patient-Reported Outcomes Measurement Information System short forms: physical function, pain interference, depression, and anxiety in knee osteoarthritis. J Pain 2017;18:1096–110.

235. Mehta SP, Fulton A, Quach C, Thistle M, Toledo C, Evans NA. Measurement properties of the Lower Extremity Functional Scale: a systematic review. J Ortho Sports Phys Ther 2016;46:200–16.

236. Rolfson O, Wissig S, van Maasakkers L, Stowell C, Ackerman I, Ayers D. Defining an international standard set of outcome measures for patients with hip or knee osteoarthritis: consensus of the International Consortium for Health Outcomes Measurement Hip and Knee Osteoarthritis Working Group. Arthritis Care Res (Hoboken) 2016;68:1631–9.

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
KOOS	42	Five domains: pain, symptoms, ADLs, recreation, and QOL	Patient-completed questionnaire	Previous week for all domains except QOL (no specification)	Five-point Likert scale (0-4)	0-100	0 = worse; 100 = no problems	Population-based normative data available	Chinese, Greek, Icelandic, Spanish, Dutch, French, Saudi Arabic, Japanese, Persian, Portuguese, Russian, English, Thai, and Turkish
KOOS-PS	7	Single domain	Patient-completed questionnaire	Previous week	Five-point Likert scale (0-4)	0-28	28 = no problems	Population-based normative data available	Specific adaptations are found in French, Portuguese, and Turkish, although the KOOS has been validated as a whole in more.
IKDC Subjective Knee Evaluation	18	Three domains: symptoms, sports and ADLs, and knee function	Patient-completed questionnaire	Not specified for some; 4 weeks for some questions; and function before and after surgery	Yes/no, Likert five-point scale, and 11-point scale	0-100	100 = no limitation	Available from the general US population	Arabic, Portuguese, Chinese, Danish, Dutch, German, Greek, Italian, Korean, Romanian, Thai, and Turkish
LKS	8	Categories of limping, support, locking, instability, pain, swelling, stair climbing, and squatting	Patient-completed questionnaire	Not specified	Individual items are scored differently using individual scoring scales.	Each item has been assigned an arbitrary score on an increasing scale. The total score is the sum of each response, of a possible score of 100.	Possible score range: 0-100 (100 = no symptoms or disability)	Normative data are available with and without stratification by sex.	Arabic, Chinese, Dutch, Italian, German, Turkish, and Spanish
TAS	11	Graduated list of ADLs, recreation, and competitive sports that describes their current level of activity	Patient-completed questionnaire	Current ability	One item is selected from a list of 11 items.	0-10	Possible score range: 0-10 (higher scores represent participation in higher-level activities)	Normative data have been presented by sex and age group.	Swedish, Dutch, German, Chinese, and Iranian

(Continued)

Table 1. (*Cont'd*)

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
ARS	4	Single index pertaining to frequency of athletic activities	Patient-completed questionnaire	1 year	Each item is followed by five responses for the frequency of each functional component within the past year	0-4	Possible score range: 0-16 (16 represents more frequent participation)	Not available	Swedish and Iranian
OKS	12	Single index that can be separated into pain and function subscales	Patient-completed questionnaire	Previous 4 weeks	Five-point Likert; 1 is best and 5 is worst	12-60	Higher scores reflect poor outcomes	No	Arabic, Chinese, Finnish, German, Japanese, Korean, Persian, Portuguese, Spanish, Swedish, Thai, and Turkish
KOS-ADL	17	Single index; two sections pertaining to symptoms and functional limitations	Patient-completed questionnaire	1-2 days	Descriptive response translated to a numerical ordinal scale	0-100	Higher scores reflect fewer knee-related symptoms and functional limitations	No	Arabic, Chinese, French, German, Portuguese, Polish, Turkish, and Greek
WOMAC	24	Three subscales: pain severity during various positions or movements, severity of joint stiffness, and difficulty performing functional activities	Self-administered or interview-administered questionnaire; in person, over the telephone, or electronically via computer or mobile phone	48 hours	Likert version: 0-4; VAS and numerical rating scale versions: 0-100 or 11-box scale	Likert format: 0-20 for pain, 0-8 for stiffness, and 0-68 for physical function; VAS format: 0-500 for pain, 0-200 for stiffness, and 0-1700 for physical function	Higher scores indicate worse pain, stiffness, or physical function	Yes, for Australian population stratified by age and sex	Arabic, Bangladesh, Chinese, Finnish, German, Greek, Hebrew, Italian, Japanese, Korean, Moroccan, Nepali, Persian, Portuguese, Singapore, Spanish, Swedish, Thai, and Turkish
PROMIS-PF	10	Five questions on limitation of activities and five questions on difficulty of ADLs	Patient-completed questionnaire; paper format and computer-adaptive testing options	No specific recall period	Likert scoring (1-5)	Raw score ranges 10-50. Scaled score 13.5 to 61.9.	50 or 61.9 representing optimal function; T-score rescales the raw score into a standardized score with a mean of 50 and an SD of 10	Yes	English, Spanish, Danish, Dutch, French, German, Hungarian, Italian, Polish, Russian, Chinese, and Ukrainian

* ADL = activity of daily living; ARS = Activity Rating Scale; IKDC = International Knee Documentation Committee; KOOS = Knee Injury and Osteoarthritis Outcome Score; KOOS-PS = KOOS Physical Function Short Form; KOS-ADL = Knee Outcome Survey Activities of Daily Living Scale; LKS = Lysholm Knee Scoring Scale; OKS = Oxford Knee Score; PROMIS-PF = Patient-Reported Outcomes Measurement Information System; QOL = quality of life; TAS = Tegner Activity Score; VAS = visual analog scale; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

Table 2. Psychometrics*

Measure	Floor and Ceiling Effects	Reliability	Validity	Responsiveness	MCIDs	Generalizability	Used in RCTs
KOOS	Little or no floor or ceiling effects for knee injury or moderate OA; floor effects for severe OA; ceiling effects following TKA	Adequate internal consistency; adequate test-retest reliability	Good content and face validity; superior convergent and divergent construct validity	Very responsive to change for OA; limited research for other injuries	MCID and moderate improvement is 8.0 and 15.6 for QOL	May be useful for a variety of conditions and populations, but its intent was OA-focused, and results may be less than optimum in other conditions. Recreational subscale may not be appropriate for some groups	Yes
KOOS-PS	None found in one study of knee OA	Adequate internal consistency; adequate test-retest reliability in mild to moderate OA; may be less adequate for severe OA	Good content and face validity are assumed because items were taken from the KOOS. Has convergent and divergent construct validity; good correlation with physical domains of other measures	Moderate to large effect sizes	Only one study showed the MCID for nonoperated OA to be 12	Use was intended for knee OA, and studies largely focus on this without many studies on other conditions	Yes
IKDC Subjective Knee Evaluation	None	Superior reliability compared with other measurements; adequate internal consistency and adequate test-retest reliability for knee injuries and some mixed pathologies	No patient contribution weakens content validity. Good face validity; high convergent and divergent construct validity	Adequate response to change, particularly in injuries as well as the pediatric population	Has been shown to be anywhere from 8.8 to 15.6	Generalizable to most populations, including pediatric populations; strongest psychometric properties in injury-related conditions	Yes
LKS	Little to no floor or ceiling effects	Test-retest reliability is adequate for use in groups with knee injuries but is less than adequate for groups with mixed knee pathologies. Reliability may be inadequate for use in individuals	The Lysholm scale has been reported as having face validity, but content validity from the patient's perspective cannot be assumed. Two studies have reported evidence of convergent and divergent construct validity	When comparing responsiveness following autologous chondrocyte implantation, the Lysholm and IKDC were the most sensitive to detecting changes when compared with MCKRS, KOOS, and SF-36	The MCID has not been calculated in any patient population	Generalizability has not been reported	Yes
TAS	Information on missing data and floor/ceiling effects is not available	Adequate test-retest reliability for groups with knee injuries and knee OA, although reliability is less than adequate for use in individuals	Content validity cannot be assumed, but evidence for convergent and divergent construct validity has been provided	Has not been reported	The MCID has not been determined	The TAS was found to be reliable in both adult and pediatric population	Yes

(Continued)

Table 2. (*Cont'd*)

Measure	Floor and Ceiling Effects	Reliability	Validity	Responsiveness	MCIDs	Generalizability	Used in RCTs
ARS	Studies consistently report no floor or ceiling effects in those with knee injury or OA.	One study has evaluated the test-retest reliability, finding adequate reliability for use in groups and individuals. It is also reliable in patients 14-18 years old.	Face, construct, and content validity have been demonstrated	Has not been reported	MCID has not been reported	The ARS was found to be reliable in both adult and pediatric population ages 14-18.	No
OKS	None prior to TKR; ceiling effect increases postoperatively; controversial postoperative floor effect	Adequate internal consistency; adequate test-retest reliability	Has content validity and internal and external validity postoperatively; correlates well with other knee-specific and general health questionnaires	Good sensitivity and responsiveness to change; large ES postoperatively; predicts revision TKR and range of motion after TKR	MDC90 and MIC after TKR: 4.15 and 9.22; nonoperated knee OA MIC: total 7.1, pain 17.3, and function 10.6	Useful in OA before and after TKR; likely applicable to multiple knee conditions	No
KOS-ADL	No floor effects; ceiling effects present after TKR	Adequate internal consistency and test-retest reliability	Good face validity; may lack content validity; correlates well with other knee-specific scales; has convergent and divergent construct validity	Good responsiveness for a variety of knee pathologies in various stages of treatment but may decrease over time	MCID after PT for knee OA: 2.2 at 2 mo and 5 at 12 mo	Useful for a variety of knee pathologies	No
WOMAC	Floor and ceiling effects exist for the stiffness subscale pre- and postoperatively, and ceiling effects exist postoperatively for all subscales.	Stiffness and function subscales have adequate internal consistency; may lower than adequate for pain subscale; mixed findings for test-retest reliability for all subscales.	Has face and content validity; good construct validity because of strong correlations with other measures; has convergent and divergent validity	High responsiveness in knee OA and chondral defects postoperatively; large ES for all subscales postoperatively, smaller ES for nonoperative management for knee OA and chondral defects	MCID after TKR: pain 22.9-36, symptoms 14.4-21.4, and function 19-33	Useful in OA and TKR; likely applicable to multiple knee conditions	Yes
PROMIS -PF	None for meniscal injuries, patellofemoral malalignment, multiligamentous injuries, and chondral disease; ceiling effects are noted in ACL injuries but below 15% cutoff.	Few studies have demonstrated reliability, although one showed high reliability in RA.	Unidimensional with high content and construct validity and adequate face validity	PROMIS assessments collectively have been shown to be responsive. PROMIS-PF is comparable with KOOS, SF-36 in ACLs and has high responsiveness in OA.	MCID has not been fully determined for knee-specific conditions. One study showed 2.45-21.55 in orthopedic patients with OA	Can be used among a variety of conditions, but measure is not specific to the knee, and further psychometric analysis is necessary.	Yes

* ACL = anterior cruciate ligament; ARS = Activity Rating Scale; ES = effect size; IKDC = International Knee Documentation Committee; KOOS = Knee Injury and Osteoarthritis Outcome Score; KOOS-PS = KOOS Physical Function Short Form; KOS-ADL = Knee Outcome Survey Activities of Daily Living Scale; LKS = Lysholm Knee Scoring Scale; MDC90 = minimum detectable change with a 90% confidence interval; MCID = minimum clinically important difference; MIC = minimal important change; OA = osteoarthritis; OKS = Oxford Knee Score; PROMIS = Patient-Reported Outcomes Measurement Information System; PROMIS-PF = PROMIS Physical Function; PT = physical therapy; RCT = randomized controlled trial; SF-36 = Short Form 36; TAS = Tegner Activity Score; TKA = total knee arthroplasty; TKR = total knee replacement; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.