DR. JOEL J GAGNIER (Orcid ID : 0000-0002-3162-3935)

Article type : Special Article

Measures of adult knee function

International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Knee Outcome Survey Activities of Daily Living Scale (KOS-ADL), Lysholm Knee Scoring Scale, Oxford Knee Score (OKS), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), Activity Rating Scale (ARS), and Tegner Activity Score (TAS), Patient Reported Outcome Measurement Information Systems- Physical Function (PROMIS-PF)

Michael McHugh[1], Erin Droy[1], Stefano Muscatelli[1], Joel J Gagnier[1-2]

1. Department of Orthopaedic Surgery, Michigan Medicine, University of Michigan, Ann Arbor, MI, USA.

2. Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA.

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

INTRODUCTION

Patient reported outcome measures (PROMs) are important to fulfill both clinical and research purposes with regards to assessing knee function in patient with a variety of knee conditions associated with injury, osteoarthritis or rheumatological disorders. For inclusion into this review, measures of knee function were required to be pertinent to rheumatology, orthopaedics, 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, and activity level. A 2011 review of 9 tools was published and focused on much of these issues as they related to rheumatology and orthopaedic surgery.

Based on the aforementioned criterion and the goals of this review, we used the same 9 measures developed specifically for patient reported knee function and perceptions: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Knee Outcome Survey Activities of Daily Living Scale (KOS-ADL), Lysholm Knee Scoring Scale, Oxford Knee Score (OKS), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), Activity Rating Scale (ARS), and Tegner Activity Score (TAS). 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 Patient-Reported Outcomes Measurement Information System Physical Function (PROMIS-PF) measure based on 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 of each patient reported outcome is displayed in table 2. Floor and ceiling effects were considered to be absent if no participants scored the bottom or top score, respectively, and acceptable if <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 alpha 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 measures ability to detect change over a period of time or intervention. Minimum clinically important difference is the amount of change of 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 studies. 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 (IKDC) SUBJECTIVE KNEE EVALUATION FORM

Description

Purpose. To detect 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 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. 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. 18 (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 5-point Likert scales; and items 2, 3, and 10 use 11-point numerical rating scales.

Recall period for items. Not specified for items 1, 3, 5, 7, 8, and 9; 4 weeks for items 2, 4, and 6. Function prior to knee injury for item 10a and current function for 10b.

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

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

Practical Application

Method of administration. 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 (i.e., 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 are summed to give a total score (excluding item 10a). The total score is calculated as (sum of items)/(maximum possible score) 100, to give a total score 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 where patients have up to 2 missing values (i.e., responses have been provided for at least 16 items), the total score is calculated as (sum of completed items)/ (maximum possible sum of completed items) 100.

Score interpretation. Possible score range 0–100, where 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. 10 minutes to complete (6). It uses simple language that is suitable for patients.

Administrative burden. 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. Available in English, Arabic, Brazilian Portuguese, traditional Chinese (Taiwan, Hong Kong), simplified Chinese (China, Singapore), Czech, Dutch, French, German, Greek, Italian, Japanese, Korean, Norwegian, Polish, Spanish, Swedish, Thai, 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, considering questions from the Standard Knee Evaluation Form, the MODEMS Lower Limb Instrument, and the Activities of Daily Living 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, items were selected by the IKDC, a committee of international orthopedic surgeons (2). Following development, validation and implementation of the IKDC, a pediatric form was developed, the Pedi-IKDC, which has been tested for psychometric properties and normative data as well as is electronic use (20-22).

Floor and ceiling effects. Studies consistently report no floor or ceiling effects (i.e., no participants scored 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 also has 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 Chineses 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 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 means that content validity cannot necessarily be assumed.

Construct validity. There are consistent reports of high convergent and divergent construct validity, with the IKDC 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 score to be highly correlated with the Cincinnati Knee Rating System, pain visual analog scale, Oxford 12 Questionnaire,

Western Ontario and McMaster Universities Osteoarthritis Index, Lysholm score, and SF-36 physical component, physical function, and bodily pain subscales (8, 13, 29).

Responsiveness. The IKDC has been shown to be adequately responsive (24). In a study comparing responsiveness of the IKDC versus KOOS for ACL injuries, the IKDC was found to be adequately responsive and the KOOS was not (30). The same finding was found in a Chinese study comparing IKDC versus 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 to 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 standard error of the measure between 3.2 and 5.6. Few studies have shown the minimal important changes. One study shows the minimal important change to be 10.9 for meniscal injuries (31). Another showed it to be 9.8 in the Chinese population (25) and another showed 12.0 for pediatric populations (12). The minimum clinically important difference has been reported to be 6.3 at 6 months and 16.7 at 12 months following cartilage repair (32), and 11.5–20.5 (range 6–28 months) in those who have undergone various surgical procedures for mixed (various) knee pathologies (33). The patient-acceptable symptom state has not been determined.

Critical Appraisal of Overall Value to the Rheumatology Community

Strengths. At face value, the domains covered by the IKDC 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 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 3 of the items may be a problem for some patients. The use of 1 aggregate score to represent symptoms, activities, and function may mask deficits in 1 domain. Psychometric testing is lacking for patients with knee osteoarthritis as an isolated group, as well as responsiveness following nonsurgical management, highlighting areas for future studies.

Clinical usability. The IKDC 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 in research for a variety of knee conditions. As some versions of the IKDC 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 (KOOS)

Description

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

Content. 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 activities of daily living (ADL); 4) difficulty experienced with sport and recreational activities; and 5) knee-related quality of life (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. 42 items across 5 subscales.

Response options/scale. All items are rated on a 5-point Likert scale (0–4), specific to each item.

Recall period for items. Previous week for pain, symptoms, ADL, and sport/recreation subscales. Not defined for QOL subscale.

Cost To Use. Free of Charge. Cost of distribution, collection and data storage not assessed.

How to Obtain. Available with associated documentation at www.koos.nu.

Practical Application

Method of administration. Patient-completed, in-person questionnaire. Can be administered on paper, tablet or computer.

Scoring. Scoring sheets (manual and computer spreadsheets) are provided on the web site. Each item is scored from 0–4. The 5 dimensions are scored separately as the sum of all corresponding items.. Scores are then transformed to a 0–100 scale (percentage of total possible score achieved)(34). If a mark is placed outside a box, the closest box is chosen. If 2 boxes are marked, that which indicates more severe problems is chosen. One or 2 missing values within a subscale are substituted with the average value for that subscale. If 2 items are missing, the response is considered invalid and a subscale score is not calculated.

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

Respondent time to complete. The KOOS takes 10 minutes to complete (34). It uses simple language and similar 1-word responses for each item.

Administrative burden. Approximately 5 minutes to score, if using the scoring spreadsheet. Can be automatically calculated with any time of data management software

Translations/adaptations.

Arabic (Egypt), Arabic (Saudi Arabia), Austria-German, Bengali (India), Czech, Chinese (Hong Kong), Chinese (Singapore), 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, Portuguese (Brazil), Polish, Romanian, Russian, Singapore, English, Slovakian, Slovenian, Spanish, Spanish (US), Spanish (Peru), Swedish, Tamil (India), Telugu (India), Thai, Turkish, Ukrainian, Urdu (India), Vietnamese, Welsh, 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 Arabic (47), Hong Kong (48), Japanese (49), Persian (50), Portuguese (51), Russian, Singapore English (40), Thai (52), and Turkish (53) translations.

Psychometric Information

Floor and ceiling effects. Studies consistently report no 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 lowest score), and ceiling effects have been reported for the pain (15–22%), sport/recreation

(16%), and QOL (17%) subscales 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, while the symptom and QOL subscales have had reports of lower as well as adequate internal consistency (Table 1). In patients with knee OA, the ADL, sport/recreation, and QOL subscales have adequate internal consistency, while the pain and symptoms subscales have reports of lower as well as 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, while 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 5 subscales, the minimal detectable change ranges from 6–12 for knee injuries and from 13.4–21.1 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/ Rec). The SDC was greater for older adults and those with knee OA than for younger and ACL cohorts (58). The standard error of the measure is reported to be lower for knee injuries than for OA.

Validity. *Face and content validity.* As well as exhibiting face validity, the direct involvement of patients with knee conditions in the development of the KOOS facilitates content validity (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 Short Form 36 (SF-36) that measure similar constructs (e.g., ADL with physical function, sport/ recreation with physical function, 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, not the 3 subscales that were based on the WOMAC (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 split by age group and knee condition for pain, symptoms, ADL, sports/recreation and QoL(58). They found that further testing was necessary for the short form as well as needed 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 on 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, while the sport/recreation and QOL subscales show moderate effects. Larger effect sizes are found following TKR than with non-operative treatment (58). Large effect sizes are consistently reported on all subscales 3–12 months after TKR, but without increasing effect sizes 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 minimum clinically important difference (MCID) of the KOOS short form and KOOS QoL has been reported in one study (62). MCID and moderate improvement estimates for 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 well as 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 (e.g., exercise therapy is likely to have more impact on ADL and sport/recreation, while pharmacology may impact more on pain and symptoms) and ensures content validity in groups of different ages and functional activity levels (e.g., the sport/ recreation subscale is more important in patients with a high physical activity level, while 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 where 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 3 WOMAC subscales facilitates 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 (KOOS-PS)

Description

Purpose. Patients' opinions about the difficulties they experience with physical activity due to their knee problems.

Content. Measure of physical function derived from the activities of daily living and sport/recreation subscales of the KOOS (63). Patients rate the degree of difficulty they have experienced over the previous week due to their knee pain, with respect to: 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. 7 items.

Response options/scale. All items are scored on a 5-point Likert scale (none, mild, moderate, severe, extreme) scored from 0–4.

Recall period for items. Previous week.

Cost to use. Free to use. 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. Patient-completed questionnaire. Can be completed in paper form or electronic form.

Scoring. Each question is scored from 0–4. The raw score is the sum of the 7 items. The interval score from 0–100 is obtained using a conversion chart (63). No instructions on how to handle missing values.

Score interpretation. Possible raw score range: 0–28. Scores are then transformed to a score from 0–100, where 0 no difficulty.

Normative values. Not available.

Respondent burden. Based on findings for the KOOS, no more than 2 minutes to complete. Uses simple language and the same 1-word responses for each of the 7 items. As the items relate to everyday tasks, it is not considered that they would have an emotional impact on the individual.

Administrative burden. Less than 5 minutes to score, using the conversion table provided (63). Training is not necessary, as the questionnaire and scoring instructions are self-explanatory.

Translations/adaptations. Available in Arabic, Chinese, Danish, Dutch, English, French, German, Hindi, Italian, Korean, Norwegian, Polish, Portuguese, Russian, Spanish, Swedish, and Turkish. Can easily be compiled by extracting the 7 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 Western Ontario and McMaster Universities Osteoarthritis Index (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 anterior cruciate ligament repair, as well as those scheduled to undergo TKR (63).

Acceptability. Rates of missing data have not been reported. Findings of 1 study indicate no floor or ceiling effects when used in patients with knee OA (i.e., no patients had 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 standard error of the measure have not been reported.

Validity. *Face and content validity.* As 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 Short Form 36 (SF-36) physical function, role physical, and bodily pain subscales; WOMAC function subscale (excluding KOOS-PS items); and Osteoarthritis Knee and Hip Quality of Life questionnaire (OAKHQOL) physical activity domain (64, 65, 67). Conversely, lower correlations have been reported with KOOS pain, symptoms, and quality of life subscales; SF-36 mental health subscales; mental health questionnaires (e.g., Profile of Mood States, Hospital Anxiety and Depression Scale); and OAKHQOL social support (64, 65, 67). One study found that, in patients with knee OA, 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 intraarticular hyaluronic acid injection (Table 2). The KOOS-PS is also able to discriminate groups of patients based on use of walking aids (65). When compared directly to 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 minimum clinically important difference (MCID) for patients undergoing non-operative treatment for OA to be 12 (70). For 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 utilized Rasch analysis in its development. Its inclusion of only 7 items facilitates use with short measures of other dimensions, such as pain visual analog scales, and makes it ideal for those for which long questionnaires may be onerous (e.g., 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, utilizing the Rasch analysis, data suggests 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 (KOS-ADL)

Description

Purpose. To determine symptoms and functional limitation in usual daily activities caused by various knee pathologies (71).

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

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

Content or domains. Single index with 2 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, 10 for function), but a 14-item version (6 for symptoms, 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 0–5, with the exception of item 9 (0–3) and item 10 (0–2) in the 17-item questionnaire.

Recall period for items. 1–2 days.

Cost to use. Free

How to obtain. Presented in full as an appendix in the original publication (71).

Practical Application

Method of administration. 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 to a percentage score by dividing by the maximum total possible score and multiplying by 100 (71,76).

Missing values. While 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 score range 0–100, where 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 as 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). There has been shown to be high test-retest reliability 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, since 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 Lysholm Knee Scoring Scale (71), WOMAC subscales (74), and global assessment of function (71). Higher correlations with the physical than 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 patient acceptable symptom state has not been reported.

Minimally important differences. Among patients with PFPS, the minimum clinically important difference 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 minimum clinically important difference 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 due to 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 as compared to 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 are utilizing.

LYSHOLM KNEE SCORING SCALE (LKS)

Description

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

Intended populations/conditions. Patients with knee ligament injury and anteromedial, anterolateral, combined anteromedial/anterolateral, posterolateral rotatory, or straight posterior instability (88).

Version. 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 8 items: 1) limp; 2) support; 3) stair climbing; 4) squatting; 5) walking, running, and jumping; and 6) thigh atrophy (88). The revised scale also includes 8 items: 1) limp, 2) support, 3) locking, 4) instability, 5) pain, 6) swelling, 7) stair climbing, and 8) squatting (89).

Number of items. 8 items.

Response options/scale. Individual items are scored differently, using individual scoring scales. The revised scale modified the original scoring slightly: 1) limp (0, 3, 5), 2) support (0, 2, 5), 3) locking (0, 2, 6, 10, 15), 4) instability (0, 5, 10, 15, 20, 25), 5) pain (0, 5, 10, 15, 20, 25), 6) swelling (0, 2, 6, 10), 7) stair climbing (0, 2, 6, 10), and 8) squatting (0, 2, 4, 5) (89).

Recall period for items. Not specified.

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

How to obtain. 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). While significantly lower scores have been found for questionnaires versus interview administration, suggesting interview bias (91), 1 study reported a high level of agreement between patients and physiotherapists using a modified version of the Lysholm scale (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) and electronically delivered questionnaires are indeed reliable modes of administration, with the perceived advantage of fostering multi-center collaborations and potential for more accurate comparisons of outcomes between patient groups (93,94).

Scoring. Each possible response to each of the 8 items has been assigned an arbitrary score on an increasing scale. The total score is the sum of each response to the 8 items, of a possible score of 100. Computer scoring is not necessary.

Missing values. No instructions provided.

Score interpretation. Possible score range: 0–100, where 100 no symptoms or disability. Scores are categorized as excellent (95–100), good (84–94), fair (65–83), and poor (<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 (i.e., patient completed versus clinician administered). The Lysholm scale 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 (i.e., clinician administered) would ensure adequate explanation of such terms, although this may vary between clinicians. As the items relate to everyday tasks, it is not considered that they would have an emotional impact on the individual.

Administrative burden. Less than 5 minutes to score. Training is not necessary, as 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, ligamentous, and meniscal injuries (97). Cross-cultural adaptations specifically for ACL injuries have been translated and validated in the Chinese and Dutch languages (98, 99) An Italian version demonstrates equivalence to the English version for assessing patellofemoral pathology (100). A German translation was found to be valid and reliable in assessing patients following total knee arthroplasty (101). Turkish and Spanish adaptations have also been accepted for use in assessing ligamentous pathology (102).

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 (103). The authors added the item for instability, as 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 were used to compare the original scale to 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 Lysholm scale.

Acceptability. Rates of missing data have not been reported. There are consistent reports of no floor or ceiling effects (i.e., 15% of patients score the lowest or highest score, respectively) (75, 90, 104-107).

Reliability. The Lysholm scale 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 as between 8.9 and 10.1 for knee injuries, while the standard error of the measure 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 Lysholm scale has been reported as having face validity, as evaluated by 5 orthopedic surgeons with sports medicine experience (75). Because the items in the Lysholm scale 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, International Knee Documentation Committee Subjective Knee Evaluation Form,

Fulkerson and Kujala scores, and Western Ontario and McMaster Universities Osteoarthritis Index (106-108). Two studies have reported evidence of convergent and divergent construct validity, finding the Lysholm score to correlate more highly with the Short Form 12 and Short Form 36 physical components than mental components (75, 90). The Lysholm score was shown to satisfy the Rasch model after 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 minimum clinically important difference (MCID) and patient-acceptable symptom state (PASS) have not been calculated in any patient population. Specifically, when comparing responsiveness following autologous chondrocyte implantation (ACI),The Lysholm and IKDC were the most sensitive to detecting changes when compared with MCKRS, KOOS, and SF-36 (109) .

Critical Appraisal of Overall Value to the Rheumatology Community

Strengths. The Lysholm scale 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 Lysholm scale assesses everyday activities as opposed to higher functional activity, delayed return to sport has little impact on the LKS. Therefore, LKS may be ideal for assessing short term outcomes, or outcomes in patients not intending to return to a specific sport (109).

Caveats and cautions. Content validity cannot be assumed, as the items included in the Lysholm scale were surgeon derived. The Lysholm scale 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 Lysholm scale in published studies. The MCID and PASS are lacking in psychometric analysis.

Clinical usability. Minimal administrative and respondent burden makes the Lysholm scale attractive for clinical use. The lack of floor and ceiling effects across different knee conditions suggests that the Lysholm scale 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 Lysholm scale is reliable for use in research on ligamentous injuries of the knee, chondral injuries, and patellar dislocation. The use of Lysholm and IKDC together has proven to represent a responsive combination for efficiently evaluating treatment effects following autologous chondrocyte implantation (109). It is important that researchers consistently utilize 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. Lack of known MCID is a weakness.

TEGNER ACTIVITY SCORE (TAS) Description

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

Intended populations/conditions. Intended for use in conjunction with the Lysholm Knee Scoring Scale, originally in patients with anterior cruciate ligament (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. Graduated list of activities of daily living, recreation, and competitive sports. The patient selects 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 to participation in national and international elite competitive sports (89). Activity levels 6–10 can only be achieved if the person participates in recreational or competitive sport.

Recall period for items. Current ability.

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

Practical Application

Method of administration. Originally established as an in-person, clinician-administered tool (110), but has been used more recently as a patient-completed questionnaire (90, 111).

Scoring. A score of 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 score range: 0–10. Higher scores represent participation in higher-level activities.

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

Respondent burden. Reported to take mean SD 3.3 0.6 minutes to complete in those who have undergone total knee replacement (112). 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 to their current activity. Degree of difficulty (measured on a visual analog scale) has been reported to increase with age (r 0.25, *P* 0.03) (112).

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

Translations/adaptations. Available in English. Cross-cultural translations are now validated for use in the Swedish, Dutch, German, Chinese, and Iranian populations for use in ACL injuries, and a German translation has been validated for use in the TKA population (99,101,113-115). 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 found to be reliable methods of administration (93).

Acceptability. Studies consistently report no floor or ceiling effects in those with knee injury or OA (i.e., 15% scored lowest or highest score, respectively)
(90, 101 , 104 , 112).

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, while the standard error of the measure ranges from 0.4–0.64.

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 and other knee injuries. However, as 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, 101, 112). The TAS has also shown significant correlations with the International Knee Documentation Committee Subjective
Knee Evaluation Form, Knee Society Score function score, Western Ontario and McMaster Universities Osteoarthritis Index pain and function subscales, and Oxford Knee Score (90, 101, 104, 112).

Generalizability. The Tegner Score was found to be reliable in both adult and pediatric populations (116).

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 minimum clinically important difference (MCID) and patient-acceptable symptom state 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 patient-reported outcomes that were developed to consider the influence of activity level on other symptoms, such as pain alleviation when aggravating activities are avoided. The Tegner score was found to have clear benefits over other PROMs, and is the preferred PROM for ACL injuries in the UK (117) .

Caveats and cautions. The TAS was originally intended and developed for patients with ACL injury as an adjunct to the Lysholm scale, 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 (118).

Clinical usability. Clinicians should note that its reliability 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 (OKS)

Description

Purpose. Brief questionnaire for patients undergoing total knee replacement (TKR) that reflected the patient's assessment of their knee-related health status and benefits of treatment (119).

Intended populations/conditions. Patients undergoing TKR.

Version. A new version was proposed on the basis that some surgeons believed that the scoring of the original version was non-intuitive (i.e., lower scores represented better outcome, higher scores represented worse outcome), where the original 12 items are used but the scoring is different (120).

Content or domains. 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 (121).

Number of items. 12 items.

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

Recall period for items. Previous 4 weeks.

Cost to use. Free.

How to obtain. The original version can be found in its original publication (119). The modified version is freely available online (http://www.orthopaedicscore.com/scorepages/oxford_knee_score.html) (120).

Practical Application

Method of administration. Patient-completed questionnaire.

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

Missing values. No instructions provided.

Score interpretation. In the original version, the total score ranges from 12–60 (119), while in the modified version the total score ranges from 0–48 (120). Higher scores in the original version reflect poor outcome and lower scores reflect better outcomes. In the modified version, this is reversed.

Normative values. Not available.

Respondent time to complete. Reported to involve minimal respondent burden (119). It takes approximately 5–10 minutes to complete the questionnaire. No training or assistance is required since the questions are self-explanatory.

Administrative burden. Scoring is simple and quick (119). Calculation of the total score takes 1–5 minutes. No training is necessary.

Translations/adaptations. Translated and validated in many languages, including Arabic (122, 123), Chinese (124-126), Finnish (127), German (128), Japanese (129), Korean (130), Persian (131, 132), Portugese (133), Spanish (134), Swedish (135), and Thai (136), Turkish (138).

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 post-operatively, with no ceiling effect (138). Furthermore, the ceiling effect has been shown to increase from 6 months to 2 years post operatively from TKR, while no floor effect was found at either time point (139). Another study found no floor or ceiling effects in patients with knee OA (134).

Reliability. The OKS has adequate internal consistency across multiple languages (119, 124, 128, 129, 135, 136) (Table 1). The original study reported adequate test–retest reliability for use in groups and individuals (119). There test-retest reliability was confirmed in patients with OA being managed non operatively (140).

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 Western Ontario and McMaster Universities Osteoarthritis Index, American Knee Society Score, Knee Outcome Survey Activities of Daily Living Scale, and pain and physical function components of the Short Form 36 and Health Assessment Questionnaire (119, 141). Internal and external validity is adequate post-operatively (138). Convergent and divergent construct validity is demonstrated by higher correlations with the Short Form 12 physical than mental component (74). Convergent and divergent validity has also been confirmed when the pain and function subscales of the OKS has been separately compared to other outcome scores (140). The OKS has been shown to fit Rasch models following rescoring of some items (142), and removal of items for limp and kneeling (143).

Responsiveness. The OKS demonstrates good sensitivity and responsiveness to change (Table 2). Responsiveness is consistent whether using raw OKS data or after Rasch analysis (139). Large effect sizes have been reported 6–12 months after TKR (119, 144) and 1 year after high tibial osteotomy (145). The OKS has also been found to be a good predictor of revision TKR within 6 months (146), and is also a predictor of range of motion after TKR (147). The effect size is larger in patients who report positive changes in their knee symptoms over time compared to patients who report a negative progression in symptoms (134).

Minimally important differences. The minimum detectable change (MDC95) after high tibial osteotomy was reported to be 8.29 (145). The minimum detectable change (MDC90) and minimal important change (MIC) 6 months after TKR were found to be 4.15 and 9.22, respectively (148). In patients with knee OA, the MDC90 after 3 months of non-operative management was 6 and the minimum important difference was 6.4 (140). 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. Patient acceptable symptom state 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 due to 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 pre-operative health status in regards to the OKS with good consistency (149), even up to 1 year post-operatively (150).

Critical Appraisal of Overall Value to the Rheumatology Community

Strengths. The OKS is a self-administered questionnaire developed to measure outcome following TKR. Due to simplicity and ease of administering, it has been used widely, especially in the UK, 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) followup of TKR compared to physician administered instruments, such as the American Knee Society Score, as reported by 1 study (151). 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 (e.g., 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 1 domain, particularly given that only 1 of the 12 items relates solely to pain. If separated into pain and function subscales, must be aware of the complex interaction between pain and function and therefore patient interpretation of the questions. The floor and ceiling effects post-operatively make post-operative 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 (152).

WESTERN ONTARIO AND MCMASTER UNIVERSITIES OSTEOARTHRITIS INDEX (WOMAC)

Description

Purpose. To assess the course of disease or response to treatment in patients with knee or hip osteoarthritis (OA) (153, 154).

Intended populations/conditions. Patients with knee and hip OA (153,154).

Version. Initially developed in 1982, the WOMAC has undergone multiple revisions (most recent version 3.1). It is available in 5-point Likert, 100-mm visual analog scale (VAS), and 11-box numerical rating scales (155, 156). Reduced and modified versions of the WOMAC have been validated but are not endorsed on the WOMAC web site (157-159, 160).

Content or domains. 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. 24 items.

Response options/scale. In the Likert version, each item offers 5 responses: "none" scored as 0, "mild" as 1, "moderate" as 2, "severe" as 3, and "extreme" as 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" (153,154).

Recall period for items. 48 hours.

Cost to use. Not free, cost depends on research project.

How to obtain. Available from Professor Nicholas Bellamy (Australia, e-mail: 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 to http://www.womac.com.

Practical Application

Method of administration. 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 (154, 161-164). Electronic and paper questionnaires show high agreement (165).

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 range for possible subscale scores in the Likert format are: pain (0–20; 5 items each scored 0–4), stiffness (2 items, 0–8), and physical function (17 items, 0–68). In the VAS

format, the ranges for the 3 subscale scores are: pain, 0–500; stiffness, 0–200; and physical function, 0–1,700 (153,154).

Missing values. If 2 or more pain items, both stiffness items, and 4 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 (153).

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

Normative values. Australian population-based normative data have been reported, stratified by age and sex (166).

Respondent time to complete. 5–10 minutes to complete.

Administrative burden. Approximately 5 minutes to score. Training is not necessary.

Translations/adaptations. WOMAC version 3.1 is available in over 100 languages (155), and has validated language translations for Arabic (167), Arabic reduced (168), Bangladesh (169), Chinese (170, Symonds 2015 171), Finnish (172), German (173), Greek (174), Hebrew (175), Italian (176), Japanese (177), Korean (178), Moroccan (179), Nepali (180) Persian (131), Portuguese-Brazil (181) Singapore (182), Spanish (183), Swedish (184, 185), Thai (186), and Turkish (187,188).

Psychometric Information

Floor and ceiling effects. Reports of floor and ceiling effects have differed between studies (74, 170, 186, 188, 189). The stiffness subscale has been reported as having floor and ceiling effects prior to intervention (74, 170, 188), as well as up to 1 year postoperatively from TKR (190). Ceiling effects have been reported by various studies for all subscales 6 months and 2 years after TKR (74, 189).

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 slightly lower than adequate. 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 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 (191). Additionally, the physical function subscale may have a stronger association with pain than performance (192,

193) The minimal detectable change and standard error of the measure 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*. Since the WOMAC was developed with extensive input from patients with OA, as well as input from academic rheumatologists and epidemiologists experienced in 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 (e.g., Short Form 36 [SF-36] physical subscales, pain/handicap VAS) suggest convergent construct validity (170, 175, 176, 179, 187, 188, 194, 195), while lower correlations with measures such as the SF-36 mental subscales indicate divergent construct validity (170, 176, 187, 188, 195). Convergent validity was also demonstrated with strong correlation with the 30 second chair stand and 50 foot timed walk tests (196). 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 (197). While 1 study found the pain subscale to demonstrate good item separation and unidimensionality in patients with knee or hip OA (198), a subsequent study found that a reduced pain subscale (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 (197). 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 (198). Similarly, Davis et al (197) 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 (199). Hip abductor and knee extensor strength are not correlated with WOMAC function subscale (200)

Responsiveness. The WOMAC appears to be responsive to change following surgical and nonsurgical interventions for knee OA and chondral defects (Table 2, 201). A recent study has confirmed the high responsiveness in patients undergoing TKR with a mean change in score of 29 at 3 months post operatively (202). In particular, the physical function domain has been suggested to be the best choice for detecting changes over time compared to other measures (203). However, in patients undergoing exercise therapy, the total WOMAC was found to be more responsive than the physical function subscale (204). In patients with knee OA, large effect sizes are consistently reported on all 3 subscales up to 2 years post-TKR. This was recently confirmed at one 1 year postoperatively, however the stiffness subscale has a smaller effect size than pain or function (205). Furthermore, effect size decreases over time up to 2 year after TKR (190). Following exercise intervention, the stiffness subscale shows small effect sizes at 2 weeks compared to 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 3 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, 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 6 and 12 months postoperatively, while 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 (206). The reduced WOMAC has been shown to have similar responsiveness to the original WOMAC (207).

Minimally important differences. The minimum clinically important difference has been calculated for TKR (up to 2 years postoperatively; range for pain 22.9–36, range for symptoms 14.4–21.4, range for function 19–33) and nonsteroidal antiinflammatory use (4 weeks; function 9.1). At 3 months post operatively, the MCID was determined to be 10.21 (202). The patient-acceptable symptom state has been determined to be 31.0 (95% confidence interval 29.4–32.9) for the function subscale in people with knee OA (208). The minimum important change has also been determined for the short WOMAC, 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 (209).

Generalizability. The WOMAC has been mainly used for OA and TKR, however it has been used in other knee pathologies. Due to the consistent reliability, validity, and responsiveness that is found in the literature it can be inferred that it 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 (210). The effect of patellofemoral overstuffing on clinical outcomes was also investigated with the WOMAC as the primary outcome (211). The WOMAC was a primary outcome measure in a trial investigating tanezumab for hip and knee arthritis (212). 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 (213). A trial of aqueous extract of Terminalia chebula fruit as a dietary supplement in overweight, healthy adults used the modified WOMAC as a primary outcome (214). Pain progression evaluated using the WOMAC was associated with radiographic and MRI evaluation of cartilage loss in symptomatic knee osteoarthritis (215).

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 (e.g., stair climbing, taking a bath) may result in missing data. Content validity is not ensured for more physically active patients since 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 compared to 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 means that the pain and function subscales are able to monitor deterioration in condition over time, while 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 due to 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. Extensive use of the WOMAC in previous research facilitates comparison of new findings.

ACTIVITY RATING SCALE (ARS) Descriptive

Purpose. Developed as a short, simple, knee-specific questionnaire to evaluate the activity level of patients with various knee disorders who participate in different sports. Intended to provide data on an athlete's highest activity level within the past year (i.e., at a time when they were most active) (216).

Intended populations/conditions. Various knee conditions, including ligament, meniscus, and chondral injury; patellofemoral pain; osteochondritis dissecans; trabecular fracture; and iliotibial band syndrome (216).

Version. No modifications to the original version.

Content. Single index pertaining to frequency of athletic activities: 1) running, 2) cutting, 3) decelerating, and 4) pivoting.

Number of items. 4 items.

Response options/scale. Each item is followed by 5 responses for the frequency of each functional component within the past year.

Recall period for items. 1 year.

Endorsements. None.

Examples of use. Conditions: anterior cruciate ligament (ACL) injury, cartilage injury, and knee osteoarthritis. Interventions: ACL reconstruction, autologous chondrocyte implantation, microfracture, high tibial osteotomy, and total knee replacement.

Practical Application

How to obtain. The ARS can be found as an appendix in the original publication (216).

Method of administration. Patient-completed questionnaire, either paper or electronically administered, with particularly high rates of agreement for the Activity Rating Scale (ARS) (217). ARS not yet been validated for interview administration (telephone, in person).

Scoring. Each item is scored from 0–4, where 0 "less than 1 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 4 items (216). *Missing values.* No specific instructions for handling missing values.

Score interpretation. The total possible score range is 0–16, where 16 represents more frequent participation. *Normative values.* Not available.

Respondent burden. Approximately 1 minute to complete. Respondent burden was intentionally minimized through the inclusion of only 4 items (216).

Administrative burden. Less than 5 minutes to score. No training is required.

Translations/adaptations. A Cross-cultural adaptation has been conducted for the Swedish translation (218), and a Persian version has been translated and validated specifically for ACL injuries (219).

Psychometric Information

Method of development. Items were selected by literature review, expert opinion (orthopedic surgeons who specialized in sports medicine, physical therapists, and athletic trainers), and 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 4, as agreed by the panel of clinicians, were retained in the final version (216).

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 (216) (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. Final item selection also involved the opinion of clinicians to ensure face validity (216).

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 Tegner Activity Score, Cincinnati Knee Ligament Score, and Daniel Score, suggesting good convergent construct validity (216).

Generalizability. Previously, the ARS had only been validated for adult use. However, a study published in 2015 found ARS to be reliable in patients younger than the age of 18 in knee injuries, with decreasing reliability in patients younger than 14. Test-retest data confirmed its reliability in all but 1 of the questions in the age >14<18 cohort. While the questionnaire may prove useful in this pediatric population, its usefulness may be limited by the significant ceiling effect observed, as more than half of the patients had maximum scores of 16 (50.6%) (220).

Ability to detect change. The responsiveness, minimum clinically important difference, and patient-acceptable symptom state 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. As it assesses 4 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 (e.g., work tasks).

Caveats and cautions. Since 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, as 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. Or 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 re-injury (221). Therefore, the utility of an activity rating scale is maximized as an adjunct to scales that assess other domains of knee function (222).

Other activities such as swimming and jumping cannot be evaluated by this scale. Furthermore, since the ARS does not focus on current ability, but on baseline activity frequency perhaps prior to injury, the validity of the instrument depends on the subject'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 (216).

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. As the scale measures the highest level of activity over the past year, without taking into account time of injury, it may be more suited for within-subject study designs, rather than comparing ratings between subjects.

PATIENT REPORTED OUTCOME MEASUREMENT INFORMATION SYSTEMS- PHYSICAL FUNCTION (PROMIS-PF)

Description

Purpose. 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, back), as well as instrumental activities of daily living, such as running errands (**223**).

Content. The Patient-Reported Outcome Measurement Information Systems (PROMIS), funded by the NIH, 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. (**224**). This tool is unique as it is useful in various disciplines in measuring physical, mental, and social health in individuals with chronic conditions (**223**). 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 (**223**).

Number of items. Form 10a has 10 items. The first 5 focus on the degree to which the patient's health limits their activities: vigorous, walking, climbing stairs, carrying groceries, bending or kneeling. The second 5 focus on difficulty in carrying out ADLs: vacuuming or yard work, dressing, shampoo hair, wash and dry body, and use the toilet.

Recall period for items. There is no specification of 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. 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 are available as well as integrated data collection tools through HealthMeasures (http://www.healthmeasures.net/resource-center/data-collection-tools).

Scoring. Scoring. Creators of the PROMIS intended the measurement to be scores according to response pattern scoring with item-level calibrations using the Health Measures Scoring Service (33TU33TUhttps://www.assessmentcenter.net/ac_scoringservice). However, there is also a table to be used. Each question has five response options, Likert, ranging 1-5. 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 10-50 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 a standard deviation (SD) of 10. Therefore, a person with a T-score of 40 is one SD below the mean.

Respondent time to complete. 5 Minutes

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

Translations and adaptations. PROMIS physical function 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, multi-ligamentous injuries, and chondral disease (**225, 226, 228**). In a study of 204 patients, PROMIS-PF was found to have no floor effect, while one patient scored the highest possible score (**227**). When compared to the KOOS-ADL, KOOS- sport and SF-36 PF, 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, well below the 15% cutoff (**229**).

Reliability. Few studies have demonstrated reliability in the PROMIS-PF form. One study showed high reliability in patients with rheumatoid arthritis (**230**). Hung and colleagues, when attempting to validate the lower

extremity physical function computer adaptive test based on PROMIS PF items found the items to demonstrate high reliability (**231**).

Validity. Hung and colleagues, 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 (**231**). Another study by the same group showed adequate face validity as well as construct validity (**232**). Good construct validity of the PROMIS-PF form in patients with rheumatoid arthritis has also been shown (**230, 233**). Content validity was further shown in patients with patients with Tenosynovial Giant Cell Tumors of the knee (**234**). Strong validity of comparisons for the PROMIS-PF items was shown in patients with different musculoskeletal disorders namely chronic pain, rheumatoid arthritis and osteoarthritis and there is high correlation with SF-36 scores (**227**, **235**). Good convergence has been found between the PROMIS-PF, KOOS and IKDC scales (**236**).

Responsiveness. The PROMIS assessments collectively have been shown to be very responsive to change (**237**). When compared to the KOOS-ADL, KOOS- sport and SF-36 PF, the PROMIS-PF showed equal responsiveness and excellent utility in the postoperative ACL course (**229**). The PROMIS-PF was specifically found to have high responsiveness to patients with OA, while pain, depression and anxiety PROMIS forms only have moderate responsiveness (**238**). It has also been shown to compare well with disease specific scales in regards to knee arthroscopy patients (**236**).

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 a variety of musculoskeletal and rheumatologically conditions. Further, with regards to the knee, it is comparable in psychometric properties to 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 to 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. Easy to use in the research setting and has been shown to be comparable to other scales that are specific to knee conditions. May allow for comparing similarities in physical function changes between patients of varying conditions.

SUMMARY / RECOMMENDATIONS

We reviewed nine of 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. Further, many tools have also been cross-culturally translated into multiple languages and adapted where needed. While other measures may be useful (e.g., 239), 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 Physical Function. When seeking to use knee measures it might be useful to refer to core outcome sets or minimum standard sets of outcomes (e.g, 240).

REFERENCES

1. Streiner DL, Norman GR. Health measurement scales: a practical guideto their development and use. 4th ed. Oxford: Oxford University Press; 2008.
2. Irrgang JJ, Anderson AF, Boland AL, Harner CD, Kurosaka M, NeyretP, 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 ligamentinjuries 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, Khamis Mohamed, et al. "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)." *SICOT-J* 5 (2019).

1107 8. Metsavaht L, Leporace G, Riberto M, de Mello Sposito MM, BatistaLA. Translation and cross-cultural
1108adaptation of the Brazilian version of the International Knee Documentation Committee subjective knee form:
1109validity and reproducibility. Am J Sports Med 2010;38:1894–9.
1110 9. Fu SN, Chan YH. Translation and validation of Chinese version ofInternational Knee Documentation
1111Committee subjective knee form. Disabil Rehabil 2011;33:1186–9.
1112 10. Jia, Zhen-yu, et al. "Translation and validation of the Simplified Chinese version of International Knee
1113Documentation Committee Subjective Knee Form." *Archives of orthopaedic and trauma surgery* 138.10 (2018):
11141433-1441.
1115 11. Fu, Siu-Ngor, and Yat-Hing Chan. "Translation and validation of Chinese version of International Knee
1116Documentation Committee subjective knee form." *Disability and rehabilitation* 33.13-14 (2011): 1186-1189.
1117 12. Jacobsen, J. S., et al. "Reproducibility and responsiveness of a D anish P edi-IKDC subjective knee form
1118for children with knee disorders." *Scandinavian journal of medicine & science in sports* 26.12 (2016): 1408-1414.
1119 13. Haverkamp D, Sierevelt IN, Breugem SJ, Lohuis K, Blankevoort L, vanDijk CN. Translation and validation
1120of the Dutch version of the International Knee Documentation Committee subjective knee form. Am J Sports Med
11212006;34:1680–4.
1122 14. Kümmel, Danica, et al. "Measurement properties of the German version of the IKDC subjective knee
1123form (IKDC-SKF)." *Journal of patient-reported outcomes* 2.1 (2018): 31.
1124 15. Koumantakis, George A., et al. "Cross-cultural adaptation and validation of the international knee
1125documentation committee subjective knee form in Greek." *Journal of Orthopaedics and Traumatology* 17.2 (2016):
1126123-129.
1127 16. Kim, Jin Goo, et al. "Translation and validation of the Korean version of the international knee
1128documentation committee subjective knee form." *Knee surgery & related research* 25.3 (2013): 106.
1129 17. Todor, Adrian, et al. "Cross-cultural adaptation and validation of the Romanian International Knee
1130Documentation Committee—subjective knee form." *PeerJ* 8 (2020): e8448.
1131 18. Lertwanich P, Praphruetkit T, Keyurapan E, Lamsam C, Kulthanan T.Validity and reliability of Thai version
1132of the International Knee Documentation Committee subjective knee form. J Med Assoc Thai 2008;91:1218–25.
1133 19. Çelik, Derya, et al. "Translation and cross-cultural adaptation of the international knee documentation
1134committee subjective knee form into Turkish." *journal of orthopaedic & sports physical therapy* 44.11 (2014): 899-
1135909.
1136 20. Kocher, Mininder S., et al. "Reliability, validity, and responsiveness of a modified International Knee
1137Documentation Committee Subjective Knee Form (Pedi-IKDC) in children with knee disorders." The American journal
1138of sports medicine 39.5 (2011): 933-939.

1139 21. Nasreddine, Adam Y., et al. "The pediatric international knee documentation committee (pedi-ikdc) 1140subjective knee evaluation form: normative data." *The American journal of sports medicine* 45.3 (2017): 527-534.
1141 22. Sabatino, M. J., et al. "An electronic patient-reported outcomes measurement system in paediatric 1142orthopaedics." *Journal of children's orthopaedics* 13.4 (2019): 431-437.
1143 23. Crawford K, Briggs KK, Rodkey WG, Steadman JR. Reliability, validity, and responsiveness of the IKDC 1144score for meniscus injuries of the knee. Arthroscopy 2007;23:839–44.
1145 24. Grevnerts, Hanna Tigerstrand, Caroline B. Terwee, and Joanna Kvist. "The measurement properties of 1146the IKDC-subjective knee form." *Knee Surgery, Sports Traumatology, Arthroscopy* 23.12 (2015): 3698-3706.
1147 25. Huang, Chien-Chih, et al. "Comparing the Chinese versions of two knee-specific questionnaires (IKDC and 1148KOOS): reliability, validity, and responsiveness." *Health and quality of life outcomes* 15.1 (2017): 238
1149 26. van de Graaf, Victor A., et al. "Reliability and validity of the IKDC, KOOS, and WOMAC for patients with 1150meniscal injuries." *The American journal of sports medicine* 42.6 (2014): 1408-1416.
1151 27. Higgins LD, Taylor MK, Park D, Ghodadra N, Marchant M, PietrobonR, et al. Reliability and validity of the 1152International Knee Documentation Committee (IKDC) subjective knee form. Joint Bone Spine 2007; 74:594–9.
1153 28. van der Velden, Charlotte A., et al. "Pedi-IKDC or KOOS-child: which questionnaire should be used in 1154children with knee disorders?." *BMC musculoskeletal disorders* 20.1 (2019): 240.
1155 29. Agel J, LaPrade RF. Assessment of differences between the modifiedCincinnati and International Knee 1156Documentation Committee patient outcome scores: a prospective study. Am J Sports Med 2009;37: 2151–7.
1157 30. van Meer, Belle L., et al. "Knee injury and Osteoarthritis Outcome Score or International Knee 1158Documentation Committee Subjective Knee Form: which questionnaire is most useful to monitor patients with an 1159anterior cruciate ligament rupture in the short term?." *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 116029.4 (2013): 701-715.
1161 31. Noorduyn, Julia CA, et al. "Responsiveness and minimal important change of the IKDC of middle-aged 1162and older patients with a meniscal tear." *The American journal of sports medicine* 47.2 (2019): 364-371.
1163 32. Greco NJ, Anderson AF, Mann BJ, Cole BJ, Farr J, Nissen CW, et al.Responsiveness of the International 1164Knee Documentation Committee subjective knee form in comparison to the Western Ontario and McMaster 1165Universities Osteoarthritis Index, modified Cincinnati Knee Rating System, and Short Form 36 in patients with focal 1166articular cartilage defects. Am J Sports Med 2010;38:891–902.
1167 33. Irrgang JJ, Anderson AF, Boland AL, Harner CD, Neyret P, RichmondJC, et al. Responsiveness of the 1168International Knee Documentation Committee subjective knee form. Am J Sports Med 2006;34:1567–73.
1169 34. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. KneeInjury and Osteoarthritis Outcome Score 1170(KOOS): development of a self-administered outcome measure. J Orthop Sports Phys Ther 1998; 28:88–96.

1171 35. Paradowski PT, Bergman S, Sunden-Lundius A, Lohmander LS, RoosEM. Knee complaints vary with age
1172and gender in the adult population: population-based reference data for the Knee injury and Osteoarthritis Outcome
1173Score (KOOS). BMC Musculoskelet Disord 2006;7:38.
1174 36. Roos EM, Roos HP, Ekdahl C, Lohmander LS. Knee injury and Osteoarthritis Outcome Score (KOOS):
1175validation of a Swedish version. Scand J Med Sci Sports 1998;8:439–48.
1176 **37.** Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis OutcomeScore (KOOS): validation and
1177comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes 2003;1:17.
1178 38. Zhang, Qi-Hao, et al. "Reliability, Validity, and Responsiveness of the Chinese Version of the Knee Injury
1179and Osteoarthritis Outcome Score (KOOS) in Patients with Anterior Cruciate Ligament Reconstruction in Mainland
1180China." Zeitschrift für Orthopädie und Unfallchirurgie 157.01 (2019): 42-47.
1181 39. Cheung, Roy TH, Shirley PC Ngai, and Kevin KW Ho. "Chinese adaptation and validation of the Knee
1182Injury and Osteoarthritis Outcome Score (KOOS) in patients with knee osteoarthritis." *Rheumatology*
1183*international* 36.10 (2016): 1449-1454.
1184 40. Xie F, Li SC, Roos EM, Fong KY, Lo NN, Yeo SJ, et al. Cross-culturaladaptation and validation of Singapore
1185English and Chinese versions of the Knee injury and Osteoarthritis Outcome Score (KOOS) in Asians with knee
1186osteoarthritis in Singapore. Osteoarthritis Cartilage 2006;14:1098–103.
1187 41. Moutzouri, Maria, et al. "Cross-cultural translation and validation of the Greek version of the Knee Injury
1188and Osteoarthritis Outcome Score (KOOS) in patients with total knee replacement." *Disability and*
1189*rehabilitation* 37.16 (2015): 1477-1483.
1190 42. Briem, Kristin. "Reliability, validity and responsiveness of the Icelandic version of the knee injury and
1191osteoarthritis outcome score (KOOS)." *Laeknabladid* 98.7-8 (2012): 403-407.
1192 43. Vaquero, Javier, et al. "Reliability, validity and responsiveness of the Spanish version of the Knee Injury
1193and Osteoarthritis Outcome Score (KOOS) in patients with chondral lesion of the knee." Knee Surgery, Sports
1194Traumatology, Arthroscopy 22.1 (2014): 104-108.
1195 44. Lizaur-Utilla, Alejandro, et al. "Validation of the Spanish Version of the Knee Injury and Osteoarthritis
1196Outcome Score (KOOS) for Elderly Patients With Total Knee Replacement." *Journal of Orthopaedic Research®* 37.10
1197(2019): 2157-2162.
1198 45. De Groot IB, Favejee MM, Reijman M, Verhaar JA, Terwee CB. TheDutch version of the Knee Injury and
1199Osteoarthritis Outcome Score: a validation study. Health Qual Life Outcomes 2008;6:16.
1200 46. Ornetti P, Parratte S, Gossec L, Tavernier C, Argenson JN, Roos EM,et al. Cross-cultural adaptation and
1201validation of the French version of the Knee injury and Osteoarthritis Outcome Score (KOOS) in knee osteoarthritis
1202patients. Osteoarthritis Cartilage 2008;16:423–8.

1203 47. Alfadhel, Saud A., et al. "Cross-cultural adaptation and validation of the Saudi Arabic version of the Knee
1204Injury and Osteoarthritis Outcome Score (KOOS)." *Rheumatology international* 38.8 (2018): 1547-1555.
1205 48. Cross-Cultural Adaptation and Validation of the Hong Kong Version of the Knee Injury and Osteoarthritis
1206Outcome Score (HK-KOOS) for Patients with Knee Osteoarthritis.
1207 49. Lyman, Stephen, et al. "Development and validation of a culturally relevant Japanese KOOS." *Journal of*
1208*Orthopaedic Science* 24.3 (2019): 514-520.
1209 50. Salavati M, Mazaheri M, Negahban H, Sohani SM, Ebrahimian MR,Ebrahimi I, et al. Validation of a
1210Persian-version of Knee injury and Osteoarthritis Outcome Score (KOOS) in Iranians with knee injuries. Osteoarthritis
1211Cartilage 2008;16:1178–82.
1212 51. Goncalves RS, Cabri J, Pinheiro JP, Ferreira PL. Cross-cultural adaptation and validation of the Portuguese
1213version of the Knee injury and Osteoarthritis Outcome Score (KOOS). Osteoarthritis Cartilage 2009; 17:1156–62.
1214 52. Chaipinyo K. Test-retest reliability and construct validity of Thaiversion of Knee Osteoarthritis Outcome
1215Score (KOOS). Thai J Phys Ther 2009;31:67–76.
1216 53. Paker N, Bugdayci D, Sabirli F, Ozel S, Ersoy S. Knee Injury andOsteoarthritis Outcome Score: reliability
1217and validation of the Turkish version. Turkiye Klinikleri J Med Sci 2007;27:350–6.
1218 **54.** Bekkers JE, de Windt TS, Raijmakers NJ, Dhert WJ, Saris DB. Validation of the Knee Injury and
1219Osteoarthritis Outcome Score (KOOS) for the treatment of focal cartilage lesions. Osteoarthritis Cartilage 2009;
122017:1434–9
1221 55. Steinhoff, Amy K., and William D. Bugbee. "Knee Injury and Osteoarthritis Outcome Score has higher
1222responsiveness and lower ceiling effect than Knee Society Function Score after total knee arthroplasty." *Knee*
1223*Surgery, Sports Traumatology, Arthroscopy* 24.8 (2016): 2627-2633.
1224 56. Örtqvist, Maria, et al. "Development of the knee injury and osteoarthritis outcome score for children
1225(KOOS-Child) comprehensibility and content validity." *Acta orthopaedica* 83.6 (2012): 666-673.
1226 57. Örtqvist, Maria, et al. "Psychometric properties of the Knee injury and Osteoarthritis Outcome Score for
1227Children (KOOS-Child) in children with knee disorders." *British journal of sports medicine* 48.19 (2014): 1437-1446.
1228 58. Collins, N. J., et al. "Knee Injury and Osteoarthritis Outcome Score (KOOS): systematic review and meta-
1229analysis of measurement properties." *Osteoarthritis and Cartilage* 24.8 (2016): 1317-1329.
1230 59. Salavati M, Akhbari B, Mohammadi F, Mazaheri M, Khorrami M.Knee injury and Osteoarthritis Outcome
1231Score (KOOS): reliability and validity in competitive athletes after anterior cruciate ligament reconstruction.
1232Osteoarthritis Cartilage 2011;19:406–10.
1233 60. Comins J, Brodersen J, Krogsgaard M, Beyer N. Rasch analysis of theKnee injury and Osteoarthritis
1234Outcome Score (KOOS): a statistical re-evaluation. Scand J Med Sci Sports 2008;18:336–45.

1235 61. Peer, Maria A., and Judith Lane. "The Knee Injury and Osteoarthritis Outcome Score (KOOS): a review of
1236its psychometric properties in people undergoing total knee arthroplasty." *Journal of Orthopaedic & Sports Physical*
1237*Therapy* 43.1 (2013): 20-28.
1238 62. Singh, Jasvinder A., et al. "Reliability and clinically important improvement thresholds for osteoarthritis
1239pain and function scales: a multicenter study." The Journal of rheumatology 41.3 (2014): 509-515.
1240 63. Perruccio AV, Stefan Lohmander L, Canizares M, Tennant A, HawkerGA, Conaghan PG, et al. The
1241development of a short measure of physical function for knee OA KOOS-Physical Function Shortform (KOOS-PS): an
1242OARSI/OMERACT initiative. Osteoarthritis Cartilage 2008;16:542–50.
1243 64. Ornetti P, Perruccio AV, Roos EM, Lohmander LS, Davis AM,Maillefert JF. Psychometric properties of the
1244French translation of the reduced KOOS and HOOS (KOOS-PS and HOOS-PS). Osteoarthritis Cartilage 2009;17:1604–
12458.
1246 65. Goncalves RS, Cabri J, Pinheiro JP, Ferreira PL, Gil J. Reliability,validity and responsiveness of the
1247Portuguese version of the Knee injury and Osteoarthritis Outcome Score-Physical Function Shortform (KOOS-PS).
1248Osteoarthritis Cartilage 2010;18:372–6.
1249 66. Gul, Ebru Demir, Ozlem Yilmaz, and Hatice Bodur. "Reliability and validity of the Turkish version of the
1250knee injury and osteoarthritis outcome score-physical function short-form (KOOS-PS)." Journal of back and
1251musculoskeletal rehabilitation 26.4 (2013): 461-466.
1252 67. Davis AM, Perruccio AV, Canizares M, Hawker GA, Roos EM, Maillefert JF, et al. Comparative, validity
1253and responsiveness of the HOOS-PS and KOOS-PS to the WOMAC physical function subscale in total joint
1254replacement for osteoarthritis. Osteoarthritis Cartilage 2009;17:843–7.
1255 68. Franchignoni, Franco, et al. "Rasch analysis of the 22 knee injury and osteoarthritis outcome score–
1256physical function items in Italian patients with knee osteoarthritis." Archives of physical medicine and rehabilitation
125794.3 (2013): 480-487.
1258 69. Mahler, E. A. M., et al. "Responsiveness of four patient-reported outcome measures to assess physical
1259function in patients with knee osteoarthritis." Scandinavian journal of rheumatology 45.6 (2016): 518-527.
1260 70. Harris, Kristina K., et al. "Extending the use of PROMs in the NHS—using the Oxford Knee Score in
1261patients undergoing non-operative management for knee osteoarthritis: a validation study." *BMJ open* 3.8 (2013):
1262e003365.
1263 71. Irrgang JJ, Snyder-Mackler L, Wainner RS, Fu FH, Harner CD. Development of a patient-reported measure
1264of function of the knee. J Bone Joint Surg Am 1998;80:1132–45.
1265 72. Marx R. Knee rating scales. Arthroscopy 2003;19:1103–8.
1266 73. Piva SR, Gil AB, Moore CG, Fitzgerald GK. Responsiveness of the activities of daily living scale of the knee
1267outcome survey and numeric pain rating scale in patients with patellofemoral pain. J Rehabil Med 2009;41:129–35.

1268 74. Impellizzeri F, Mannion A, Leunig M, Bizzini M, Naal F. Comparison of the reliability, responsiveness, and
1269construct validity of 4 different questionnaires for evaluating outcomes after total knee arthroplasty. J Arthroplasty
12702010. E-pub ahead of print.
1271 75. Marx RG, Jones EC, Allen AA, Altchek DW, O'Brien SJ, Rodeo SA,et al. Reliability, validity, and
1272responsiveness of four knee outcome scales for athletic patients. J Bone Joint Surg Am 2001;83-A:1459–69.
1273 76. Irrgang J. Development of a health related quality of life instrument toassess physical function related to
1274pathology and impairment of the knee. Pittsburgh: University of Pittsburgh; 1999.
1275 77. Algarni AD, Alrabai HM, Al-Ahaideb A, Kachanathu SJ, AlShammari SA. Arabic translation, cultural
1276adaptation, and validation study of Knee Outcome Survey: Activities of Daily Living Scale (KOS-ADLS). *Rheumatol Int*.
12772017;37(9):1585-1589. doi:10.1007/s00296-017-3776-4
1278 78. Bouzubar FF, Aljadi SH, Alotaibi NM, Irrgang JJ. Cross-cultural adaptation and validation of the Arabic
1279version of the knee outcome survey-activities for daily living scale. *Disabil Rehabil*. 2018;40(15):1817-1828.
1280doi:10.1080/09638288.2017.1313319
1281 79. Jia Z-Y, Wang W, Nian X-W, et al. Cross-cultural Adaptation and Validation of the Simplified Chinese
1282Version of the Knee Outcome Survey Activities of Daily Living Scale. *Arthroscopy*. 2016;32(10):2009-2016.
1283doi:10.1016/j.arthro.2016.01.068
1284 80. Roy J-S, Esculier J-F, Maltais DB. Translation, cross-cultural adaptation and validation of the French
1285version of the Knee Outcome Survey-Activities of Daily Living Scale. *Clin Rehabil*. 2014;28(6):614-623.
1286doi:10.1177/0269215513511342
1287 81. Bizzini M, Gorelick M. Development of a German version of the kneeoutcome survey for daily activities.
1288Arch Orthop Trauma Surg 2007; 127:781–9.
1289 82. Goncalves R, Cabri J, Pinheiro J. Cross-cultural adaptation and validation of the Portuguese version of
1290the Knee Outcome SurveyActivities of Daily Living Scale (KOS-ADLS). Clin Rheumatol 2008; 27:1445–9.
1291 83. Szczepanik M, Bejer A, Snela S, Szymczyk D, Jabłoński J, Majewska J. Polish Cross-Cultural Adaptation
1292and Validation of the Knee Outcome Survey Activities of Daily Living Scale (KOS-ADLS) in Patients Undergoing Total
1293Knee Arthroplasty. *Med Sci Monit*. 2018;24:5309-5319. doi:10.12659/MSM.908094
1294 84. Evcik D, Ay S, Ege A, Turel A, Kavuncu V. Adaptation and validationof Turkish version of the Knee
1295Outcome Survey-Activities for Daily Living Scale. Clin Orthop Relat Res 2009;467:2077–82.
1296 85. Kapreli E, Panelli G, Strimpakos N, Billis E, Zacharopoulos A, Athanasopoulos S. Cross-cultural adaptation
1297of the Greek version of the Knee Outcome Survey-Activities of Daily Living Scale (KOS-ADLS). Knee 2010.
1298 86. Esculier J-F, Roy J-S, Bouyer LJ. Psychometric evidence of self-reported questionnaires for patellofemoral
1299pain syndrome: a systematic review. *Disabil Rehabil*. 2013;35(26):2181-2190. doi:10.3109/09638288.2013.774061

1300 87. Williams VJ, Piva SR, Irrgang JJ, Crossley C, Fitzgerald GK. Comparison of reliability and responsiveness of
1301patient-reported clinical outcome measures in knee osteoarthritis rehabilitation. *J Orthop Sports Phys Ther*.
13022012;42(8):716-723. doi:10.2519/jospt.2012.4038
1303 88. Lysholm J, Gillquist J. Evaluation of knee ligament surgery resultswith special emphasis on use of a
1304scoring scale. Am J Sports Med 1982;10:150–4.
1305 89. Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. Clin Orthop Relat Res
13061985;198:43–9.
1307 90. Briggs KK, Lysholm J, Tegner Y, Rodkey WG, Kocher MS, SteadmanJR. The reliability, validity, and
1308responsiveness of the Lysholm Score and Tegner Activity Scale for anterior cruciate ligament injuries of the knee: 25
1309years later. Am J Sports Med 2009;37:890–7.
1310 91. Hoher J, Bach T, Munster A, Bouillon B, Tiling T. Does the modeof data collection change results in a
1311subjective knee score? Selfadministration versus interview. Am J Sports Med 1997;25:642–7.
1312 92. Smith HJ, Richardson JB, Tennant A. Modification and validation ofthe Lysholm Knee Scale to assess
1313articular cartilage damage. Osteoarthritis Cartilage 2009;17:53–8.
1314 93. Nguyen, Joseph, Robert Marx, Chisa Hidaka, Sean Wilson, and Stephen Lyman. 2017. “Validation of
1315Electronic Administration of Knee Surveys among ACL-Injured Patients.” *Knee Surgery, Sports Traumatology,*
1316*Arthroscopy* 25 (10): 3116–22. https://doi.org/10.1007/s00167-016-4189-8.
1317 94. Kose, Ozkan, Gokmen Deniz, Hakan Ozcan, and Ferhat Guler. 2015. “A Comparison of Telephone
1318Interview versus On-Site Completion of Lysholm Knee Score in Patients Who Underwent Arthroscopic ACL
1319Reconstruction: Are the Results Equivalent?” *European Journal of Orthopaedic Surgery and Traumatology* 25 (6):
13201069–72. https://doi.org/10.1007/s00590-015-1605-6.
1321 95. Briggs KK, Steadman JR, Hay CJ, Hines SL. Lysholm score and Tegneractivity level in individuals with
1322normal knees. Am J Sports Med 2009;37:898–901.
1323 96. Demirdjian AM, Petrie SG, Guanche CA, Thomas KA. The outcomesof two knee scoring questionnaires in
1324a normal population. Am J Sports Med 1998;26:46–51.
1325 97. Ahmed, Khamis Mohamed, Hatem G. Said, Eslam Karam Allah Ramadan, Mohamed Abd El-Radi, and
1326Maher A. El-Assal. 2019. “Arabic Translation and Validation of Three Knee Scores, Lysholm Knee Score (LKS), Oxford
1327Knee Score (OKS), and International Knee Documentation Committee Subjective Knee Form (IKDC).” *SICOT-J* 5: 6.
1328https://doi.org/10.1051/sicotj/2018054.
1329 98. Wang, W., L. Liu, X. Chang, Z. Y. Jia, J. Z. Zhao, and W. D. Xu. 2016. “Cross-Cultural Translation of the
1330Lysholm Knee Score in Chinese and Its Validation in Patients with Anterior Cruciate Ligament Injury.” *BMC*
1331*Musculoskeletal Disorders* 17 (1): 436. https://doi.org/10.1186/s12891-016-1283-5.

1332 99. Eshuis, Rienk, Gijsbertus Wilhelmus Lentjes, Yelverton Tegner, Nienke Wolterbeek, and Maurits
1333Remmelt Veen. 2016. “Dutch Translation and Cross-Cultural Adaptation of the Lysholm Score and Tegner Activity
1334Scale for Patients with Anterior Cruciate Ligament Injuries.” *Journal of Orthopaedic and Sports Physical Therapy* 46
1335(11): 976–83. https://doi.org/10.2519/jospt.2016.6566
1336 100. Cerciello, Simone, Katia Corona, Brent Joseph Morris, Enrico Visonà, Giulio Maccauro, Nicola Maffulli,
1337and Mario Ronga. 2018. “Cross-Cultural Adaptation and Validation of the Italian Versions of the Kujala, Larsen,
1338Lysholm and Fulkerson Scores in Patients with Patellofemoral Disorders.” *Journal of Orthopaedics and Traumatology*
133919 (1). https://doi.org/10.1186/s10195-018-0508-9
1340 101. Swanenburg, Jaap, Peter P. Koch, Nadja Meier, and Brigitte Wirth. 2014. “Function and Activity in
1341Patients with Knee Arthroplasty: Validity and Reliability of a German Version of the Lysholm Score and the Tegner
1342Activity Scale.” *Swiss Medical Weekly* 144 (June). https://doi.org/10.4414/smw.2014.13976.
1343 102. Arroyo-Morales, Manuel, Jose Martin-Alguacil, Mario Lozano-Lozano, Antonio I. Cuesta-Vargas, Andrés J.
1344Fernández-Fernández, Jose A. González, Yelverton Tegner, and Irene Cantarero-Villanueva. 2019. “The Lysholm
1345Score: Cross Cultural Validation and Evaluation of Psychometric Properties of the Spanish Version.” *PLoS ONE* 14 (8):
1346e0221376. https://doi.org/10.1371/journal.pone.0221376.
1347 103. Oretorp N, Gillquist J, Liljedahl SO. Long term results of surgery fornon-acute anteromedial rotatory
1348instability of the knee. Acta Orthop Scand 1979;50:329–36.
1349 104. Briggs KK, Kocher MS, Rodkey WG, Steadman JR. Reliability, validityand responsiveness of the Lysholm
1350Knee Score and the Tegner Activity Scale for patients with meniscal injury of the knee. J Bone Joint Surg Am
13512006;88:698–705.
1352 105. Heintjes EM, Bierma-Zeinstra SM, Berger MY, Koes BW. Lysholmscale and WOMAC index were
1353responsive in prospective cohort of young general practice patients. J Clin Epidemiol 2008;61:481–8.
1354 106. Kocher MS, Steadman JR, Briggs KK, Sterett WI, Hawkins RJ. Reliability, validity, and responsiveness of
1355the Lysholm knee scale for various chondral disorders of the knee. J Bone Joint Surg Am 2004;86-A:1139– 45.
1356 107. Paxton EW, Fithian DC, Stone ML, Silva P. The reliability and validityof knee-specific and general health
1357instruments in assessing acute patellar dislocation outcomes. Am J Sports Med 2003;31:487–92.
1358 108. Sgaglione NA, Del Pizzo W, Fox JM, Friedman MJ. Critical analysis ofknee ligament rating systems. Am J
1359Sports Med 1995;23:660–7.
1360 109. Howard, Jennifer S., Christian Lattermann, Johanna M. Hoch, Carl G. Mattacola, and Jennifer M. Medina
1361McKeon. 2013. “Comparing Responsiveness of Six Common Patient-Reported Outcomes to Changes Following
1362Autologous Chondrocyte Implantation: A Systematic Review and Meta-Analysis of Prospective Studies.” *Cartilage* 4
1363(2): 97–110. https://doi.org/10.1177/1947603512470684.

1364 110. Hambly K. The use of the Tegner Activity Scale for articular cartilagerepair of the knee: a systematic
1365review. Knee Surg Sports Traumatol Arthrosc 2011;19:604–14.
1366 111. Frobell RB, Roos EM, Roos HP, Ranstam J, Lohmander LS. A randomized trial of treatment for acute
1367anterior cruciate ligament tears. N Engl J Med 2010;363:331–42.
1368 112. Naal FD, Impellizzeri FM, Leunig M. Which is the best activity ratingscale for patients undergoing total
1369joint arthroplasty? Clin Orthop Relat Res 2009;467:958–65.
1370 113. Flosadottir, Vala, Ewa M. Roos, and Eva Ageberg. 2017. "Translation, Cross-Cultural Adaptation, and
1371Validation of the Activity Rating Scale for Disorders of the Knee." *Orthopaedic Journal of Sports Medicine* 5 (9).
1372https://doi.org/10.1177/2325967117729361.
1373 114. Mostafaee, Neda, Hossein Negahban, Mohammad Jafar Shaterzadeh Yazdi, Shahin Goharpey,
1374Mohammad Mehravar, and Nahid Pirayeh. 2018. "Responsiveness of a Persian Version of Knee Injury and
1375Osteoarthritis Outcome Score and Tegner Activity Scale in Athletes with Anterior Cruciate Ligament Reconstruction
1376Following Physiotherapy Treatment." *Physiotherapy Theory and Practice*, 2018.
1377https://doi.org/10.1080/09593985.2018.1548672.
1378 115. Huang, Hongshi, Dongxia Zhang, Yanfang Jiang, Jie Yang, Tao Feng, Xi Gong, Jianquan Wang, and
1379Yingfang Ao. 2016. "Translation, Validation and Cross-Cultural Adaptation of a Simplified-Chinese Version of the
1380Tegner Activity Score in Chinese Patients with Anterior Cruciate Ligament Injury." *PLoS ONE* 11 (5): e0155463.
1381https://doi.org/10.1371/journal.pone.0155463.
1382 116. Sabatino, M. J., C. V. Gans, A. J. Zynda, J. S. Chung, S. M. Miller, P. L. Wilson, C. H. Jo, and H. B. Ellis. 2019.
1383"An Electronic Patient-Reported Outcomes Measurement System in Paediatric Orthopaedics." *Journal of Children's*
1384*Orthopaedics* 13 (4): 431–37. https://doi.org/10.1302/1863-2548.13.190053.
1385 117. Letchford, Robert, Valerie Sparkes, and Robert W.M. van Deursen. 2015. "Assessing Participation in the
1386 ACL Injured Population: Selecting a Patient Reported Outcome Measure on the Basis of Measurement
1387 Properties." *Knee* 22 (3): 262–69. https://doi.org/10.1016/j.knee.2015.01.010.
1388 118. Frobell RB, Svensson E, Gothrick M, Roos EM. Self-reported activitylevel and knee function in amateur
1389football players: the influence of age, gender, history of knee injury and level of competition. Knee Surg Sports
1390Traumatol Arthrosc 2008;16:713–9.
1391 119. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on theperceptions of patients about total knee
1392replacement. J Bone Joint Surg Br 1998;80:63–9.
1393 120. Murray D, Fitzpatrick R, Rogers K, Pandit H, Beard D, Carr A, et al.The use of the Oxford hip and knee
1394scores. J Bone Joint Surg Br 2007;89:1010–4.

1395 121. Harris KK, Dawson J, Jones LD, Beard DJ, Price AJ. Extending the use of PROMs in the NHS--using the
1396Oxford Knee Score in patients undergoing non-operative management for knee osteoarthritis: a validation study.
1397*BMJ Open*. 2013;3(8):e003365. doi:10.1136/bmjopen-2013-003365
1398 122. Ahmed KM, Said HG, Ramadan EKA, Abd El-Radi M, El-Assal MA. Arabic translation and validation of
1399three knee scores, Lysholm Knee Score (LKS), Oxford Knee Score (OKS), and International Knee Documentation
1400Committee Subjective Knee Form (IKDC). *SICOT J*. 2019;5:6. doi:10.1051/sicotj/2018054
1401 123. Alghadir A, Anwer S, Iqbal ZA, Alsanawi HA. Cross-cultural adaptation, reliability and validity of the
1402Arabic version of the reduced Western Ontario and McMaster Universities Osteoarthritis index in patients with knee
1403osteoarthritis. *Disabil Rehabil*. 2016;38(7):689-694. doi:10.3109/09638288.2015.1055380
1404 124. Xie F, Li S, Lo N, Yeo S, Yang K, Yeo W, et al. Cross-cultural adaptation and validation of Singapore English
1405and Chinese Versions of the Oxford Knee Score (OKS) in knee osteoarthritis patients undergoing total knee
1406replacement. Osteoarthritis Cartilage 2007;15:1019–24.
1407 125. Cheung RTH, Ngai SPC, Ho KKW. Chinese translation and validation of the Oxford Knee Scale for patients
1408with knee osteoarthritis. *Hong Kong Physiother J*. 2017;37:46-49. doi:10.1016/j.hkpj.2017.03.002
1409 126. Lin K, Bao L, Wang J, Fujita K, Makimoto K, Liao X. Validation of the Chinese (Mandarin) Version of the
1410Oxford Knee Score in Patients with Knee Osteoarthritis. *Clin Orthop Relat Res*. 2017;475(12):2992-3004.
1411doi:10.1007/s11999-017-5495-2
1412 127. Reito A, Järvistö A, Jämsen E, et al. Translation and validation of the 12-item Oxford knee score for use in
1413Finland. *BMC Musculoskelet Disord*. 2017;18(1):74. doi:10.1186/s12891-017-1405-8
1414 128. Naal F, Impellizzeri F, Sieverding M, Loibl M, Von Knoch F, MannionA, et al. The 12-item Oxford Knee
1415Score: cross-cultural adaptation into German and assessment of its psychometric properties in patients with
1416osteoarthritis of the knee. Osteoarthritis Cartilage 2009;17:49–52.
1417 129. Takeuchi R, Sawaguchi T, Nakamura N, Ishikawa H, Saito T,Goldhahn S. Cross-cultural adaptation and
1418validation of the Oxford 12-item knee score in Japanese. Arch Orthop Trauma Surg 2011;131: 247–54.
1419 130. Eun IS, Kim OG, Kim CK, Lee HS, Lee JS. Validation of the Korean version of the Oxford Knee Score in
1420patients undergoing total knee arthroplasty. *Clin Orthop Relat Res*. 2013;471(2):600-605. doi:10.1007/s11999-012-
14212564-4
1422 131. Ebrahimzadeh MH, Makhmalbaf H, Birjandinejad A, Keshtan FG, Hoseini HA, Mazloumi SM. The Western
1423Ontario and McMaster Universities Osteoarthritis Index (WOMAC) in Persian Speaking Patients with Knee
1424Osteoarthritis. *Arch Bone Jt Surg*. 2014;2(1):57-62.
1425 132. Ebrahimzadeh MH, Makhmalbaf H, Birjandinejad A, Soltani-Moghaddas SH. Cross-cultural adaptation
1426and validation of the persian version of the oxford knee score in patients with knee osteoarthritis. *Iran J Med Sci*.
14272014;39(6):529-535.

1428 133. Gonçalves RS, Tomás AM, Martins DI. Cross-cultural adaptation and validation of the Portuguese version
1429of the Oxford Knee Score (OKS). *Knee*. 2012;19(4):344-347. doi:10.1016/j.knee.2011.04.006
1430 134. Martín-Fernández J, García-Maroto R, Sánchez-Jiménez FJ, et al. Validation of the Spanish version of the
1431Oxford knee score and assessment of its utility to characterize quality of life of patients suffering from knee
1432osteoarthritis: a multicentric study. *Health Qual Life Outcomes*. 2017;15(1):186. doi:10.1186/s12955-017-0761-2
1433 135. Dunbar M, Robertsson O, Ryd L, Lidgren L. Translation and validationof the Oxford-12 item knee score
1434for use in Sweden. Acta Orthop Scand 2000;71:268–74.
1435 136. Charoencholvanich K, Pongcharoen B. Oxford knee score and SF-36: translation & reliability for use with
1436total knee arthroscopy patients in Thailand. J Med Assoc Thai 2005;88:1194–202.
1437 137. Tuğay BU, Tuğay N, Güney H, Kınıklı Gİ, Yüksel İ, Atilla B. Oxford Knee Score: cross-cultural adaptation
1438and validation of the Turkish version in patients with osteoarthritis of the knee. *Acta Orthop Traumatol Turc*.
14392016;50(2):198-206. doi:10.3944/AOTT.2015.15.0127
1440 138. Jenny J-Y, Diesinger Y. The Oxford Knee Score: compared performance before and after knee
1441replacement. *Orthop Traumatol Surg Res*. 2012;98(4):409-412. doi:10.1016/j.otsr.2012.03.004
1442 139. Ko Y, Lo N-N, Yeo S-J, et al. Comparison of the responsiveness of the SF-36, the Oxford Knee Score, and
1443the Knee Society Clinical Rating System in patients undergoing total knee replacement. *Qual Life Res*.
14442013;22(9):2455-2459. doi:10.1007/s11136-013-0376-y
1445 140. Harris K, Dawson J, Doll H, et al. Can pain and function be distinguished in the Oxford Knee Score in a
1446meaningful way? An exploratory and confirmatory factor analysis. *Qual Life Res*. 2013;22(9):2561-2568.
1447doi:10.1007/s11136-013-0393-x
1448 141. Bream E, Charman SC, Clift B, Murray D, Black N. Relationship between patients' and clinicians'
1449assessments of health status before and after knee arthroplasty. *Qual Saf Health Care*. 2010;19(6):e6.
1450doi:10.1136/qshc.2008.031310
1451 142. Conaghan PG, Emerton M, Tennant A. Internal construct validity ofthe Oxford Knee Scale: evidence from
1452Rasch measurement. Arthritis Rheum 2007;57:1363–7.
1453 143. Ko Y, Lo N, Yeo S, Yang K, Yeo W, Chong H, et al. Rasch analysis ofthe Oxford Knee Score. Osteoarthritis
1454Cartilage 2009;17:1163–9.
1455 144. Garratt A, Brealey S, Gillespie W. Patient-assessed health instrumentsfor the knee: a structured review.
1456Rheumatology (Oxford) 2004;43: 1414–23
1457 145. Goldhahn S, Takeuchi R, Nakamura N, Nakamura R, Sawaguchi T. Responsiveness of the Knee Injury and
1458Osteoarthritis Outcome Score (KOOS) and the Oxford Knee Score (OKS) in Japanese patients with high tibial
1459osteotomy. *J Orthop Sci*. 2017;22(5):862-867. doi:10.1016/j.jos.2017.04.013

1460 146. Moonot P, Medalla G, Matthews D, Kalairajah Y, Field R. Correlationbetween the Oxford Knee and
1461American Knee Society scores at midterm follow-up. J Knee Surg 2009;22:226–30.
1462 147. Maempel JF, Clement ND, Brenkel IJ, Walmsley PJ. Range of movement correlates with the Oxford knee
1463score after total knee replacement: A prediction model and validation. *Knee*. 2016;23(3):511-516.
1464doi:10.1016/j.knee.2016.01.009
1465 148. Beard DJ, Harris K, Dawson J, et al. Meaningful changes for the Oxford hip and knee scores after joint
1466replacement surgery. *J Clin Epidemiol*. 2015;68(1):73-79. doi:10.1016/j.jclinepi.2014.08.009
1467 149. Kwong E, Neuburger J, Black N. Agreement between retrospectively and contemporaneously collected
1468patient-reported outcome measures (PROMs) in hip and knee replacement patients. *Qual Life Res*. 2018;27(7):1845-
14691854. doi:10.1007/s11136-018-1823-6
1470 150. Yeoman TFM, Clement ND, Macdonald D, Moran M. Recall of preoperative Oxford Hip and Knee Scores
1471one year after arthroplasty is an alternative and reliable technique when used for a cohort of patients. *Bone Joint*
1472*Res*. 2018;7(5):351-356. doi:10.1302/2046-3758.75.BJR-2017-0259.R1
1473 151. Medalla GA, Moonot P, Peel T, Kalairajah Y, Field RE. Cost-benefitcomparison of the Oxford Knee Score
1474and the American Knee Society Score in measuring outcome of total knee arthroplasty. J Arthroplasty 2009;24:652–
14756.
1476 152. Maempel JF, Clement ND, Brenkel IJ, Walmsley PJ. Validation of a prediction model that allows direct
1477comparison of the Oxford Knee Score and American Knee Society clinical rating system. *Bone Joint J*. 2015;97-
1478B(4):503-509. doi:10.1302/0301-620X.97B4.34867
1479 153. Bellamy N. WOMAC Osteoarthritis Index user guide. London (Ontario, Canada): University of Western
1480Ontario; 1995.
1481 154. Bellamy N. WOMAC Osteoarthritis Index user guide. Version V. Brisbane (Australia): CONROD, The
1482University of Queensland; 2002.
1483 155. WOMAC-AUSCAN-osteoarthritis global index. URL: http://www.womac.org.
1484 156. Ornetti P, Dougados M, Paternotte S, Logeart I, Gossec L. Validation ofa numerical rating scale to assess
1485functional impairment in hip and knee osteoarthritis: comparison with the WOMAC function scale. Ann Rheum Dis
14862011;70:740–6.
1487 157. Baron G, Tubach F, Ravaud P, Logeart I, Dougados M. Validation of ashort form of the Western Ontario
1488and McMaster Universities Osteoarthritis Index function subscale in hip and knee osteoarthritis. Arthritis Rheum
14892007;57:633–8.
1490 158. Whitehouse SL, Lingard EA, Katz JN, Learmonth ID. Development andtesting of a reduced WOMAC
1491function scale. J Bone Joint Surg Br 2003;85:706–11.

1492 159. Yang KG, Raijmakers NJ, Verbout AJ, Dhert WJ, Saris DB. Validationof the short-form WOMAC function
1493scale for the evaluation of osteoarthritis of the knee. J Bone Joint Surg Br 2007;89:50–6.
1494 160. Shen ZD, Yu HM, Wang JT, Shi GY, Sun Y. [Modified Western Ontario and McMaster University
1495Osteoarthritis Index Scale used in patients with knee osteoarthritis]. *Zhonghua Yi Xue Za Zhi*. 2019;99(7):537-541.
1496doi:10.3760/cma.j.issn.0376-2491.2019.07.012
1497 161. Bellamy N, Campbell J, Stevens J, Pilch L, Stewart C, Mahmood Z.Validation study of a computerized
1498version of the Western Ontario and McMaster Universities VA3.0 Osteoarthritis Index. J Rheumatol 1997;24:2413–5.
1499 162. Bellamy N, Campbell J, Hill J. A comparative study of telephone vson-site completion of the WOMAC 3.0
1500Osteoarthritis Index. J Rheumatol 2002;29:783–6.
1501 163. Bellamy N, Wilson C, Hendrikz J, Whitehouse SL, Patel B, Dennison S,et al. Osteoarthritis Index delivered
1502by mobile phone (m-WOMAC) is valid, reliable, and responsive. J Clin Epidemiol 2011;64:182–90.
1503 164. Theiler R, Speilberger J, Bischoff H, Bellamy N, Huber J, Kroesen S.Clinical evaluation of the WOMAC 3.0
1504OA Index in numeric rating scale format using a computerised touch screen version. Osteoarthritis Cartilage
15052002;10:479–81.
1506 165. Marsh JD, Bryant DM, Macdonald SJ, Naudie DDR. Patients respond similarly to paper and electronic
1507versions of the WOMAC and SF-12 following total joint arthroplasty. *J Arthroplasty*. 2014;29(4):670-673.
1508doi:10.1016/j.arth.2013.07.008
1509 166. Bellamy N, Wilson C, Hendrikz J. Population-based normative valuesfor the Western Ontario and
1510McMaster (WOMAC) Osteoarthritis Index and the Australian/Canadian (AUSCAN) hand osteoarthritis index
1511functional subscales. Inflammopharmacology 2010;18:1–8.
1512 167. Guermazi M, Poiraudeau S, Yahia M, Mezganni M, Fermanian J, HabibElleuch M, et al. Translation,
1513adaptation and validation of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) for an
1514Arab population: the Sfax modified WOMAC. Osteoarthritis Cartilage 2004;12:459–68.
1515 168. Alghadir A, Anwer S, Iqbal ZA, Alsanawi HA. Cross-cultural adaptation, reliability and validity of the
1516Arabic version of the reduced Western Ontario and McMaster Universities Osteoarthritis index in patients with knee
1517osteoarthritis. *Disabil Rehabil*. 2016;38(7):689-694. doi:10.3109/09638288.2015.1055380
1518 169. Rabbani MG, Haq SA, Bellamy N, et al. Development, linguistic and clinimetric validation of the WOMAC
1519VA3.01 Bangla for Bangladesh Index. *Rheumatol Int*. 2015;35(6):997-1003. doi:10.1007/s00296-014-3192-y
1520 170. Xie F, Li SC, Goeree R, Tarride JE, O'Reilly D, Lo NN, et al. Validationof Chinese Western Ontario and
1521McMaster Universities Osteoarthritis Index (WOMAC) in patients scheduled for total knee replacement. Qual Life
1522Res 2008;17:595–601.

1523 171. Symonds T, Hughes B, Liao S, Ang Q, Bellamy N. Validation of the Chinese Western Ontario and
1524McMaster Universities Osteoarthritis Index in Patients From Mainland China With Osteoarthritis of the Knee.
1525*Arthritis Care Res (Hoboken)*. 2015;67(11):1553-1560. doi:10.1002/acr.22631
1526 172. Soininen JV, Paavolainen PO, Gronblad MA, Kaapa EH. Validationstudy of a Finnish version of the
1527Western Ontario and McMasters University Osteoarthritis Index. Hip Int 2008;18:108–11.
1528 173. Stucki G, Meier D, Stucki S, Michel BA, Tyndall AG, Dick W, et al.Evaluation of a German version of
1529WOMAC (Western Ontario and McMaster Universities) Arthrosis Index. Z Rheumatol 1996;55:40–9. In German.
1530 174. Papathanasiou G, Stasi S, Oikonomou L, et al. Clinimetric properties of WOMAC Index in Greek knee
1531osteoarthritis patients: comparisons with both self-reported and physical performance measures. *Rheumatol Int*.
15322015;35(1):115-123. doi:10.1007/s00296-014-3043-x
1533 175. Wigler I, Neumann L, Yaron M. Validation study of a Hebrew versionof WOMAC in patients with
1534osteoarthritis of the knee. Clin Rheumatol 1999;18:402–5.
1535 176. Salaffi F, Leardini G, Canesi B, Mannoni A, Fioravanti A, Caporali R,et al. Reliability and validity of the
1536Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index in Italian patients with osteoarthritis of
1537the knee. Osteoarthritis Cartilage 2003;11:551–60.
1538 177. Hashimoto H, Hanyu T, Sledge CB, Lingard EA. Validation of a Japanese patient-derived outcome scale
1539for assessing total knee arthroplasty: comparison with Western Ontario and McMaster Universities Osteoarthritis
1540Index (WOMAC). J Orthop Sci 2003;8:288–93.
1541 178. Bae SC, Lee HS, Yun HR, Kim TH, Yoo DH, Kim SY. Cross-culturaladaptation and validation of Korean
1542Western Ontario and McMaster Universities (WOMAC) and Lequesne osteoarthritis indices for clinical research.
1543Osteoarthritis Cartilage 2001;9:746–50.
1544 179. Faik A, Benbouazza K, Amine B, Maaroufi H, Bahiri R, Lazrak N, et al.Translation and validation of
1545Moroccan Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index in knee osteoarthritis.
1546Rheumatol Int 2008;28:677–83.
1547 180. Nakarmi S, Haq SA, Vaidya B. Translation, validation and cross-cultural adaptation of the Nepali version
1548of WOMAC® LK 3.1. *Int J Rheum Dis*. 2019;22(10):1877-1883. doi:10.1111/1756-185X.13690
1549 181. Lage PTS, Machado LAC, Barreto SM, de Figueiredo RC, Telles RW. Measurement properties of
1550Portuguese-Brazil Western Ontario and McMaster Universities osteoarthritis index (WOMAC) for the assessment of
1551knee complaints in Brazilian adults: ELSA-Brasil Musculoskeletal cohort. *Rheumatol Int*. 2020;40(2):233-242.
1552doi:10.1007/s00296-019-04496-1
1553 182. Thumboo J, Chew LH, Soh CH. Validation of the Western Ontario andMcMaster University Osteoarthritis
1554Index in Asians with osteoarthritis in Singapore. Osteoarthritis Cartilage 2001;9:440–6.

1555 183. Escobar A, Quintana JM, Bilbao A, Azkarate J, Guenaga JI. Validationof the Spanish version of the
1556WOMAC questionnaire for patients with hip or knee osteoarthritis: Western Ontario and McMaster Universities
1557Osteoarthritis Index. Clin Rheumatol 2002;21:466–71.
1558 184. Roos EM, Klassbo M, Lohmander LS. WOMAC osteoarthritis index: reliability, validity, and
1559responsiveness in patients with arthroscopically assessed osteoarthritis. Western Ontario and McMaster
1560Universities. Scand J Rheumatol 1999;28:210–5.
1561 185. Soderman P, Malchau H. Validity and reliability of Swedish WOMACOsteoarthritis Index: a self-
1562administered disease-specific questionnaire (WOMAC) versus generic instruments (SF-36 and NHP). Acta Orthop
1563Scand 2000;71:39–46.
1564 186. Kuptniratsaikul V, Rattanachaiyanont M. Validation of a modifiedThai version of the Western Ontario
1565and McMaster (WOMAC) Osteoarthritis Index for knee osteoarthritis. Clin Rheumatol 2007;26: 1641–5.
1566 187. Basaran S, Guzel R, Seydaoglu G, Guler-Uysal F. Validity, reliability,and comparison of the WOMAC
1567osteoarthritis index and Lequesne algofunctional index in Turkish patients with hip or knee osteoarthritis. Clin
1568Rheumatol 2010;29:749–56.
1569 188. Tuzun EH, Eker L, Aytar A, Daskapan A, Bayramoglu M. Acceptability, reliability, validity and
1570responsiveness of the Turkish version of WOMAC osteoarthritis index. Osteoarthritis Cartilage 2005;13:28–33.
1571 189. Escobar A, Quintana JM, Bilbao A, Arostegui I, Lafuente I, VidaurretaI. Responsiveness and clinically
1572important differences for the WOMAC and SF-36 after total knee replacement. Osteoarthritis Cartilage
15732007;15:273–80.
1574 190. Giesinger K, Hamilton DF, Jost B, Holzner B, Giesinger JM. Comparative responsiveness of outcome
1575measures for total knee arthroplasty. *Osteoarthr Cartil*. 2014;22(2):184-189. doi:10.1016/j.joca.2013.11.001
1576 191. Ayala A, Bilbao A, Garcia-Perez S, Escobar A, Forjaz MJ. Scale invariance and longitudinal stability of the
1577Physical Functioning Western Ontario and MacMaster Universities Osteoarthritis Index using the Rasch model.
1578*Rheumatol Int*. 2018;38(3):473-479. doi:10.1007/s00296-017-3901-4
1579 192. White DK, Master H. Patient-Reported Measures of Physical Function in Knee Osteoarthritis. *Rheum Dis*
1580*Clin North Am*. 2016;42(2):239-252. doi:10.1016/j.rdc.2016.01.005
1581 193. Riddle DL, Stratford PW. Unilateral vs bilateral symptomatic knee osteoarthritis: associations between
1582pain intensity and function. *Rheumatology (Oxford)*. 2013;52(12):2229-2237. doi:10.1093/rheumatology/ket291
1583 194. Brazier J, Harper R, Munro J, Walters S, Snaith M. Generic and condition-specific outcome measures for
1584people with osteoarthritis of the knee. Rheumatology (Oxford) 1999;38:870–7.
1585 195. McConnell S, Kolopack P, Davis AM. The Western Ontario and McMaster Universities Osteoarthritis
1586Index (WOMAC): a review of its utility and measurement properties. Arthritis Rheum 2001;45:453–61.

1587 196. Gill SD, de Morton NA, Mc Burney H. An investigation of the validity of six measures of physical function
1588in people awaiting joint replacement surgery of the hip or knee. *Clin Rehabil*. 2012;26(10):945-951.
1589doi:10.1177/0269215511434993
1590 197. Davis A, Badley E, Beaton D, Kopec J, Wright J, Young N, et al. Raschanalysis of the Western Ontario
1591McMaster (WOMAC) Osteoarthritis Index: results from community and arthroplasty samples. J Clin Epidemiol
15922003;56:1076–83.
1593 198. Wolfe F, Kong SX. Rasch analysis of the Western Ontario McMasterquestionnaire (WOMAC) in 2,205
1594patients with osteoarthritis, rheumatoid arthritis, and fibromyalgia. Ann Rheum Dis 1999;58:563–8.
1595 199. Giesinger JM, Kuster MS, Behrend H, Giesinger K. Association of psychological status and patient-
1596reported physical outcome measures in joint arthroplasty: a lack of divergent validity. *Health Qual Life Outcomes*.
15972013;11:64. doi:10.1186/1477-7525-11-64
1598 200. Tevald MA, Murray A, Luc BA, Lai K, Sohn D, Pietrosimone B. Hip abductor strength in people with knee
1599osteoarthritis: A cross-sectional study of reliability and association with function. *Knee*. 2016;23(1):57-62.
1600doi:10.1016/j.knee.2015.06.006
1601 201. Trudeau J, Van Inwegen R, Eaton T, et al. Assessment of pain and activity using an electronic pain diary
1602and actigraphy device in a randomized, placebo-controlled crossover trial of celecoxib in osteoarthritis of the knee.
1603*Pain Pract*. 2015;15(3):247-255. doi:10.1111/papr.12167
1604 202. SooHoo NF, Li Z, Chenok KE, Bozic KJ. Responsiveness of patient reported outcome measures in total
1605joint arthroplasty patients. *J Arthroplasty*. 2015;30(2):176-191. doi:10.1016/j.arth.2014.09.026
1606 203. Mahler E, Cuperus N, Bijlsma J, et al. Responsiveness of four patient-reported outcome measures to
1607assess physical function in patients with knee osteoarthritis. *Scand J Rheumatol*. 2016;45(6):518-527.
1608doi:10.3109/03009742.2016.1140226
1609 204. French HP, Fitzpatrick M, FitzGerald O. Responsiveness of physical function outcomes following
1610physiotherapy intervention for osteoarthritis of the knee: an outcome comparison study. *Physiotherapy*.
16112011;97(4):302-308. doi:10.1016/j.physio.2010.03.002
1612 205. Bolink S a. a. N, Grimm B, Heyligers IC. Patient-reported outcome measures versus inertial performance-
1613based outcome measures: A prospective study in patients undergoing primary total knee arthroplasty. *Knee*.
16142015;22(6):618-623. doi:10.1016/j.knee.2015.04.002
1615 206. Parkes MJ, Callaghan MJ, Tive L, Lunt M, Felson DT. Responsiveness of Single versus Composite
1616Measures of Pain in Knee Osteoarthritis. *J Rheumatol*. 2018;45(9):1308-1315. doi:10.3899/jrheum.170928
1617 207. Jogi P, Spaulding SJ, Zecevic AA, Overend TJ, Kramer JF. Comparison of the original and reduced versions
1618of the Berg Balance Scale and the Western Ontario and McMaster Universities Osteoarthritis Index in patients
1619following hip or knee arthroplasty. *Physiother Can*. 2011;63(1):107-114. doi:10.3138/ptc.2009-26

1620 208. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al.Evaluation of clinically relevant
1621states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. Ann
1622Rheum Dis 2005;64:34–7.
1623 209. Abbott JH, Hobbs C, Gwynne-Jones D, MOA Trial Team. The ShortMAC: Minimum Important Change of a
1624Reduced Version of the Western Ontario and McMaster Universities Osteoarthritis Index. *J Orthop Sports Phys Ther*.
16252018;48(2):81-86. doi:10.2519/jospt.2018.7676
1626 210. Baranova IV. [The use of the functional state of the joints for the estimation of the effectiveness of the
1627application of oxygen/ozone therapy for the rehabilitative treatment of the patients suffering from knee arthritis].
1628*Vopr Kurortol Fizioter Lech Fiz Kult*. 2018;95(3):42-48. doi:10.17116/kurort201895342
1629 211. Kemp MA, Metcalfe AJ, Sayers A, Wylde V, Eldridge JD, Blom AW. Does overstuffing of the
1630patellofemoral joint in total knee arthroplasty have a significant effect on postoperative outcomes? *Knee*.
16312018;25(5):874-881. doi:10.1016/j.knee.2018.05.007
1632 212. Tive L, Bello AE, Radin D, et al. Pooled analysis of tanezumab efficacy and safety with subgroup analyses
1633of phase III clinical trials in patients with osteoarthritis pain of the knee or hip. *J Pain Res*. 2019;12:975-995.
1634doi:10.2147/JPR.S191297
1635 213. Quicke JG, Foster NE, Croft PR, Ogollah RO, Holden MA. Change in physical activity level and clinical
1636outcomes in older adults with knee pain: a secondary analysis from a randomised controlled trial. *BMC*
1637*Musculoskelet Disord*. 2018;19(1):59. doi:10.1186/s12891-018-1968-z
1638 214. Lopez HL, Habowski SM, Sandrock JE, et al. Effects of dietary supplementation with a standardized
1639aqueous extract of Terminalia chebula fruit (AyuFlex®) on joint mobility, comfort, and functional capacity in healthy
1640overweight subjects: a randomized placebo-controlled clinical trial. *BMC Complement Altern Med*. 2017;17(1):475.
1641doi:10.1186/s12906-017-1977-8
1642 215. Wirth W, Hunter DJ, Nevitt MC, et al. Predictive and concurrent validity of cartilage thickness change as
1643a marker of knee osteoarthritis progression: data from the Osteoarthritis Initiative. *Osteoarthr Cartil*.
16442017;25(12):2063-2071. doi:10.1016/j.joca.2017.08.005
1645 216. Marx R, Stump T, Jones E, Wickiewicz T, Warren R. Development andevaluation of an activity rating scale
1646for disorders of the knee. Am J Sports Med 2001;29:213–8.
1647 217. Nguyen, Joseph, Robert Marx, Chisa Hidaka, Sean Wilson, and Stephen Lyman. 2017. “Validation of
1648Electronic Administration of Knee Surveys among ACL-Injured Patients.” *Knee Surgery, Sports Traumatology,*
1649*Arthroscopy* 25 (10): 3116–22. https://doi.org/10.1007/s00167-
1650 218. Iversen, Maura D., Johan Von Heideken, Elisabeth Farmer, Jessica Rihm, Benton E. Heyworth, and
1651Mininder S. Kocher. 2016. “Validity and Comprehensibility of Physical Activity Scales for Children with Sport
1652Injuries.” *Journal of Pediatric Orthopaedics* 36 (3): 278–83. https://doi.org/10.1097/BPO.0000000000000448.

1653 219. Mostafaee, Neda, Hossein Negahban, Mohammad Jafar Shaterzadeh Yazdi, Shahin Goharpey,
1654Mohammad Mehravar, and Nahid Pirayeh. 2018. “Responsiveness of a Persian Version of Knee Injury and
1655Osteoarthritis Outcome Score and Tegner Activity Scale in Athletes with Anterior Cruciate Ligament Reconstruction
1656Following Physiotherapy Treatment.” *Physiotherapy Theory and Practice*, 2018.
1657https://doi.org/10.1080/09593985.2018.1548672.
1658 220. Shirazi, Cameron P, Heidi A Israel, and Scott G Kaar. 2016. “Is the Marx Activity Scale Reliable in Patients
1659Younger Than 18 Years?” *Sports Health* 8 (2): 145–48. https://doi.org/10.1177/1941738115600388.
1660 221. Cameron, Kenneth L, Karen Y Peck, Brandon S Thompson, Steven J Svoboda, Brett D Owens, and
1661Stephen W Marshall. 2015. “Reference Values for the Marx Activity Rating Scale in a Young Athletic Population:
1662History of Knee Ligament Injury Is Associated With Higher Scores.” *Sports Health* 7 (5): 403–8.
1663https://doi.org/10.1177/1941738115576121.
1664 222. Rick W. Knee injury outcomes measures. J Am Acad Orthop Surg2009;17:31–9.
1665 223. Cella, David, et al. "The Patient-Reported Outcomes Measurement Information System (PROMIS)
1666developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008." *Journal of*
1667*clinical epidemiology* 63.11 (2010): 1179-1194.
1668 224. SubCommittee, S. "PROMIS Instrument Development and Validation Scientific Standards."
1669 225. Gulbrandsen, Trevor R., et al. "Preoperative Performance of Patient-Reported Outcomes Measurement
1670Information System in Patients with Meniscal Root Tears." The Journal of Knee Surgery (2019).
1671 226. Carender, Christopher N., et al. "Preoperative Performance of PROMIS in Patients With Patellofemoral
1672Malalignment and Chondral Disease." *Orthopaedic journal of sports medicine* 7.7 (2019): 2325967119855001.
1673 227. Driban, Jeffrey B., et al. "Patient-Reported Outcomes Measurement Information System (PROMIS)
1674instruments among individuals with symptomatic knee osteoarthritis: a cross-sectional study of floor/ceiling effects
1675and construct validity." BMC musculoskeletal disorders 16.1 (2015): 253.
1676 228. Trasolini, Nicholas A., et al. "Performance of PROMIS Computer Adaptive Testing As Compared With
1677Established Instruments for Multiple-Ligament Knee Injuries." Orthopaedic journal of sports medicine 7.9 (2019):
16782325967119867419.
1679 229. Scott, Elizabeth J., et al. "Performance of the PROMIS in patients after anterior cruciate ligament
1680reconstruction." Orthopaedic journal of sports medicine 6.5 (2018): 2325967118774509.
1681 230. Bartlett, Susan J., et al. "Reliability and validity of selected PROMIS measures in people with rheumatoid
1682arthritis." PloS one 10.9 (2015).
1683 231. Hung, Man, et al. "A lower extremity physical function computerized adaptive testing instrument for
1684orthopaedic patients." Foot & ankle international 33.4 (2012): 326-335.

1685 232. Hung, Man, et al. "Validation of PROMIS® Physical Function computerized adaptive tests for orthopaedic
1686foot and ankle outcome research." Clinical Orthopaedics and Related Research® 471.11 (2013): 3466-3474.
1687 233. Oude Voshaar, Martijn AH, et al. "Validity and measurement precision of the PROMIS physical function
1688item bank and a content validity–driven 20-item short form in rheumatoid arthritis compared with traditional
1689measures." Rheumatology 54.12 (2015): 2221-2229.
1690 234. Gelhorn, Heather L., et al. "The measurement of physical functioning among patients with tenosynovial
1691giant cell tumor (TGCT) using the Patient-Reported Outcomes Measurement Information System (PROMIS)." Journal
1692of patient-reported outcomes 3.1 (2019): 6.
1693 235. Crins, Martine HP, et al. "Differential item functioning of the PROMIS physical function, pain
1694interference, and pain behavior item banks across patients with different musculoskeletal disorders and persons
1695from the general population." Quality of Life Research 28.5 (2019): 1231-1243.
1696 236. Kenney, Raymond J., et al. "Do patient reported outcome measurement information system (PROMIS)
1697scales demonstrate responsiveness as well as disease-specific scales in patients undergoing knee arthroscopy?." The
1698American journal of sports medicine 47.6 (20)
1699 237. Fries, James, Matthias Rose, and Eswar Krishnan. "The PROMIS of better outcome assessment:
1700responsiveness, floor and ceiling effects, and Internet administration." The Journal of rheumatology 38.8 (2011):
17011759-1764.
1702 238. Lee, Augustine C., et al. "Responsiveness and minimally important differences for 4 patient-reported
1703outcomes measurement information system short forms: physical function, pain interference, depression, and
1704anxiety in knee osteoarthritis." The Journal of Pain 18.9 (2017): 1096-1110.
1705 239. Mehta SP, et al. Measurement properties of the lower extremity functional scale: A systematic review. J
1706Ortho Sports Phys Ther. 2016;46(3): 200-16.
1707 240. Rolfson O et al. Defining an International Standard Set of Outcome Measures for Patients With Hip or
1708Knee Osteoarthritis: Consensus of the International Consortium for Health Outcomes Measurement Hip and Knee
1709Osteoarthritis Working Group. Arthritis Care Res (Hoboken). 2016;68(11):1631-1639. doi: 10.1002/acr.22868.

1710

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
Knee Injury and Osteoarthritis Outcome Score	42	5: pain, symptoms, ADls, recreation, QoL	Patient completed questionnaire	Previous week for all except QoL, no specification.	5 point Likert, 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, Turkish.
Knee Injury and Osteoarthritis Outcome Score-Physical Function	7	Single Domain	Patient completed questionnaire	Previous week	5 point Likert, 0-4	0-28	28 = no problems	Population based normative data available	Specific adaptation is found in French, Portuguese and Turkish although KOOS has been validated

									as a whole by more.
International Knee Documentation Committee Subjective Knee Evaluation	18	3:Symptoms, sports and ADLs, knee function	Patient completed questionnaire	Not specified for some, 4 weeks for some questions, and function before and after surgery	Yes/no; Likert 5 point scale; 11 point scale	0-100	100 is no limitation	Available from the general US population	Arabic, Portuguese, Chinese, Danish, Dutch, German, Greek, Italian, Korean, Romanian, Thai, and Turkish
Lysholm Knee Scoring Scale (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,	Possible score range: 0–100, where 100 = no symptoms or disability.	Normative data are available with and without stratification by sex.	Arabic, Chinese, Dutch, Italian, German, Turkish and Spanish

						of a possible score of 100.			
Tegner Activity Score (TAS)	11	Graduated list of activities of daily living, recreation, and competitive sports that described their current level of activity.	Patient completed questionnaire	Current ability	One item is selected from a list of 11.	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
Activity Rating Scale (ARS)	4	Single index pertaining to frequency of athletic activities.	Patient completed questionnaire	1-year recall period	Each item is followed by 5 responses for the frequency of each functional component within the past year.	0-4	Possible score range: 0–16, where 16 represents more frequent participation.	Not available.	Swedish and Iranian

Oxford Knee Score	12	Single index; can be separated into pain and function subscales	Patient completed questionnaire	Previous 4 weeks	5 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, Turkish
Knee Outcome Survey Activities of Daily Living Scale	17	Single index; 2 sections pertaining to symptoms and functional limitations	Patient completed questionnaire	1-2 days	Descriptive response translated to numerical ordinal scale	0-100	Higher scores reflect less knee-related symptoms and less functional limitations	No	Arabic, Chinese, French, German, Portuguese, Polish, Turkish, Greek
Western Ontario and McMaster Universities Osteoarthritis Index	24	Three subscales; pain severity during various positions or movements, severity of	Self-administered or interview administered questionnaire; in person, over the telephone, or	48 hours	Likert version 0-4. VAS and numerical rating scale versions	Likert format 0-20 pain, 0-8 stiffness, 0-68 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,

		joint stiffness, and difficulty performing functional activities	electronically via computer or mobile phone		with 0-100 or 11 box scale	VAS format pain 0-500, stiffness 0-200, physical function 0-1700			Japanese, Korean, Moroccan, Nepali, Persian, Portuguese, Singapore, Spanish, Swedish, Thai, Turkish
Patient Reported Outcome Measurement Information Systems-Physical Function	10	5 questions on limitation of activities, 5 questions on difficulty of ADLs.	Patient completed questionnaire. Paper format, computer adaptive testing options	No specific recall period	Likert scoring from 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 w/ mean of 50 and a standard deviation of 10.	Yes	English, Spanish, Danish, Dutch, French, German, Hungarian, Italian, Polish, Russian, Chinese, Ukranian.

Table 2: Psychometrics

Measure	Floor, ceiling effects	Reliability	Validity	Responsiveness	Minimally important differences (MCID)	Generalizability	Used in RCTs
Knee Injury and Osteoarthritis Outcome Score	Little or no floor or ceiling effects for knee injury or mod. OA. Floor effects for severe OA and 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
Knee Injury and Osteoarthritis Outcome Score-Physical	None found in 1 study of knee OA	Adequate internal consistency, adequate test-retest reliability in mild to moderate OA. May be less	Assume good content and face validity since items taken from KOOS. Convergent and	Moderate to large effect sizes	Only one study showed MCID for non-op OA to be 12	Use was intended for knee OA and studies largely focus on this without many	Yes

Function		adequate for severe OA	divergent construct validity, good correlation with physical domains of other measures.			studies on other conditions.	
International Knee Documentation Committee Subjective Knee Evaluation	None	Superior reliability over other measurements have been shown. Adequate internal consistency, 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 population. Strongest psychometric properties in injury related conditions.	Yes
Lysholm Knee Scoring Scale (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	The Lysholm scale has been reported as having face validity, content validity from the patient's perspective cannot be assumed. Two	When comparing responsiveness following autologous chondrocyte implantation (ACI),The Lysholm and IKDC were the most sensitive to	The MCID has not been calculated in any patient population.	Generalizability has not been reported.	Yes

		inadequate for use in individuals.	studies have reported evidence of convergent and divergent construct validity.	detecting changes when compared with MCKRS, KOOS, and SF-36.			
Tegner Activity Score (TAS)	Information on missing data and floor/ceiling effects is not available.	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.	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 Tegner Score was found to be reliable in both adult and pediatric population.	Yes
Activity Rating Scale (ARS)	Studies consistently report no floor or ceiling effects in those with knee injury or OA.	One study has evaluated the test–retest reliability of the ARS, finding adequate reliability for use in groups and individuals. Also reliable in patients 14-18.	Face, construct, and content validity have been demonstrated.	Has not been reported.	MCID has not been reported.	ARS was found to be reliable in both adult and pediatric population ages 14-18.	No

Oxford knee score	None prior to TKR; ceiling effect increases post-operatively; controversial post-op floor effect	Adequate internal consistency, adequate test-retest reliability	Content validity; internal and external validitiy post-operatively; correlates well with other knee-specific and general health questionnaires	Good sensitivity and responsiveness to change; large ES post-op; predicts revision TKR and range of motion after TKR	MDC90 and MIC after TKR: 4.15 and 9.22; non-op knee OA MIC: total 7.1, pain 17.3, function 10.6	Useful in OA before and after TKR; likely applicable to multiple knee conditions	No
Knee Outcome Survey Activities of Daily Living Scale	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; 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, 5 at 12mo	Useful for a variety of knee pathologies	No
Western Ontario and McMaster Universities Osteoarthritis Index	Floor and ceiling effects exist for the stiffness subscale pre and post-op and ceiling effects exist post-op for all	Stiffness and function subscales have adequate internal consistency, may lower than adequate for pain subscale; mixed findings for test-	Has face and content validity; Good construct validity due to strong correlations with other measures, convergent and divergent validity	High responsiveness in knee OA and chondral defects post-op; large ES all subscales post-op, smaller ES for non-op management	MCID after TKR: for pain 22.9-36, symptoms 14.4-21.4, function 19-33	Useful in OA and TKR; likely applicable to multiple knee conditions	Yes

	subscales	retest reliability for all subscales		knee OA and chondral defects			
Patient Reported Outcome Measurement Information Systems-Physical Function	None for meniscal injuries, patellofemoral malalignment, multi-ligamentous injuries and chondral disease. Ceiling effects 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. Adequate face validity.	PROMIS assessments collectively have been shown to be responsive. PROMIS-PF compared to KOOS, SF-36 in ACLs. High responsiveness in OA.	MCID has not been fully determined for knee specific conditions. One study showed 2.45 to 21.55 in orthopaedic patients w/ OA.	Can be use among a variety of conditions, but measure is not specific to the knee and further psychometric analysis is necessary.	Yes