Published in final edited form as:
Arthritis Care Res (Hoboken). 2011 November ; 63(0 11): S208–S228. doi:10.1002/acr.20632.

Measures of Knee Function:

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

NATALIE J. COLLINS[1], **DEVYANI MISRA**[2], **DAVID T. FELSON**[2], **KAY M. CROSSLEY**[1], and **EWA M. ROOS**[3]

[1]Natalie J. Collins, PhD, PT, Kay M. Crossley, PhD, PT: The University of Melbourne, Melbourne, Victoria, Australia

[2]Devyani Misra, MD, David T. Felson, MD, MPH: Boston University School of Medicine, Boston, Massachusetts

[3]Ewa M. Roos, PhD, PT: University of Southern Denmark, Odense, Denmark

INTRODUCTION

Patient-reported measures of knee function are important for the comprehensive assessment of rheumatology conditions in both clinical and research contexts. To merit inclusion in this review, measures of knee function were required to be patient reported and assess aspects considered important by adult patients with knee problems such as injury or osteoarthritis (OA). Therefore, measures used in rheumatology, orthopedics, and sports medicine were considered. Dimensions deemed to be important to patients included pain, function, quality of life, and activity level. To identify instruments fulfilling these criteria, we utilized published reviews of knee instruments (1), knee OA instruments (2), and measures for use in patellofemoral arthroplasty (3).

Based on these reviews, as well as extensive searches of more recent literature, we included the following 9 patient-reported outcomes: Activity Rating Scale, International Knee Documentation Committee Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score, Knee Injury and Osteoarthritis Outcome Score Physical

Address correspondence to Ewa M. Roos, PhD, PT, Research Unit for Musculoskeletal Function and Physiotherapy, Institute of Sports Science and Clinical Biomechanics, University of Southern Denmark, Campusvej 55, DK-5230, Odense M, Denmark. eroos@health.sdu.dk.

Dr. Roos has received speaking fees (less than $10,000) from Biomet.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published.

Function Short Form, Knee Outcome Survey Activities of Daily Living Scale, Lysholm Knee Scoring Scale, Tegner Activity Scale, Oxford Knee Score, and Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). Although the WOMAC can be applied to the hip and knee, this study contains data only applicable to the knee. Measures assessing activity level are listed separately.

Psychometric data pertaining to the reliability and responsiveness of each patient-reported outcome are shown in Tables 1 and 2. The number of psychometric reports concerning each instrument ranges from 2–27. A higher number of reports indicates a higher degree of certainty in interpretation of the psychometric properties.

Psychometric properties were based on data provided in Tables 1 and 2, and interpreted using standardized guidelines. Internal consistency was considered adequate if Cronbach's alpha was at least 0.7 (4), and test–retest (intra-rater) reliability was adequate if the intraclass correlation coefficient was at least 0.8 for groups and 0.9 for individuals (5). 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 (6,7). We defined content validity as present when there was patient involvement in the development and/or selection of items (7). Measures were deemed to have face validity if the reviewers considered that the items adequately reflected the measured construct, or if studies reported that expert panels had made a similar assessment (8). 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 (7). As there is no gold standard measure of patient-reported outcome, criterion validity is not applicable to this review. Effect sizes of <0.5 were considered small, 0.5–0.8 were considered moderate, and >0.8 were considered large (9). In this context, the minimum clinically important difference is the amount of change of a patient-reported outcome that represents a meaningful change to the patient, while the patient-acceptable symptom state is the least abnormal function score at which patients would consider themselves having acceptable function (10).

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 (11).

Intended populations/conditions: Patients with a variety of knee conditions, including ligament injuries, meniscal injuries, articular cartilage lesions, and patellofemoral pain (11).

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 (12) and revised in 1994 (13). 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. The items now have the allocated scores next to each possible response. The minimum score for each item has also been changed so that it is now 0, not 1. The scoring of the numerical rating scales for items 2 and 3 has been reversed so that 0 represents the highest level of symptoms and 10 represents the lowest level of symptoms, which is in line with the scoring of the rest of the items.

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) (11).

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.

Endorsements—International Cartilage Repair Society; European Society of Sports Traumatology, Knee Surgery, and Arthroscopy; and American Orthopaedic Society for Sports Medicine (AOSSM).

Examples of use—Conditions: knee ligament injury (anterior cruciate ligament [ACL], posterior cruciate ligament [PCL], lateral collateral ligament [LCL], medial patello-femoral ligament), meniscal tears, knee cartilage lesions, osteochondritis dissecans, and traumatic knee dislocation. Interventions: ligament reconstruction (ACL, PCL, LCL, medial patellofemoral ligament), meniscal repair, meniscectomy, microfracture, osteochondral autografts, platelet-rich plasma injections, high tibial osteotomy, and lateral release.

Practical Application

How to obtain—The most recent revision is freely available at the AOSSM web site as part of the IKDC Knee Forms (2000; www.sportsmed.org/tabs/research/ikdc.aspx). Multiple web sites have published versions of the form.

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.

Missing values: 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 values: Normative data are available from the general US population, stratified for age, sex, and current/prior knee problems (14).

Respondent burden—10 minutes to complete (15). 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, traditional Chinese (Taiwan, Hong Kong), simplified Chinese (China, Singapore), French, German, Italian, Japanese, Korean, Portuguese (Brazil), and Spanish. Cross-cultural adaptations have been conducted for the Brazilian (16), Chinese (17), Dutch (18), Italian (15), and Thai (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 (11).

Acceptability—Missing data were relatively common in testing of the final version of the form, with 57 of 590 patients failing to answer >3 items of 18 (11). Studies consistently report no floor or ceiling effects (i.e., no participants scored lowest or highest score) (11,15,16, 18,20).

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. However, test–retest reliability is slightly below adequate for individuals who fall into a broader category of knee

pathologies. 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.

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 (11,16–18,20,21). 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 (16,18,22).

Ability to detect change—In patients undergoing surgical treatment of meniscal injury, the IKDC shows large effect sizes at 1 year (Table 2). For patients who have had surgical intervention for cartilage injury, the IKDC shows moderate effect sizes at 6 months and large effect sizes at 1 year. Large effect sizes have been reported from 6–28 months following various surgical procedures conducted in a mixed cohort of knee pathologies. The minimum clinically important difference has been reported to be 6.3 at 6 months and 16.7 at 12 months following cartilage repair (23), and 11.5–20.5 (range 6–28 months) in those who have undergone various surgical procedures for mixed (various) knee pathologies (24). 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.

Caveats and cautions—Despite demonstrating face validity, the lack of patient contribution to item selection indicates that content validity cannot necessarily be assumed. 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 non-surgical 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. Test–retest reliability for those with various knee pathologies suggests that the IKDC may demonstrate inadequate reliability for the evaluation of individual patients.

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 patients' opinions about their knee and associated problems over short- and long-term followup (1 week to decades).

Intended populations/conditions: Young and middle-aged people with posttraumatic osteoarthritis (OA), as well as those with injuries that may lead to posttraumatic OA (e.g., anterior cruciate ligament [ACL], meniscal, or chondral injury) (25).

Version: The original KOOS remains unchanged, although a short form for function has been developed.

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) (25).

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.

Endorsements—International Cartilage Repair Society, American Academy of Orthopedic Surgeons, and US Food and Drug Administration.

Examples of use—Conditions: 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.

Practical Application

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

Method of administration—Patient-completed, in-person questionnaire. The KOOS has not been validated for use during an in-person or telephone interview.

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. A total score has not been validated and is not recommended. Scores are then transformed to a 0–100 scale (percentage of total possible score achieved), where 0 = extreme knee problems and 100 = no knee problems (25).

Missing values: 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 = extreme problems and 100 = no problems.

Normative values: Population-based normative data are available, stratified by age and sex (26).

Respondent burden—The KOOS takes 10 minutes to complete (25). It uses simple language and similar 1-word responses for each item. The items largely reflect signs and symptoms of their knee condition and how this affects everyday tasks, so it is not considered that they would have an emotional impact on the individual. The knee-related QOL subscale could be considered the most emotionally sensitive component, as it requires the individual to reflect on how their knee affects their QOL.

Administrative burden—Approximately 5 minutes to score, using the scoring spreadsheet. Training is not necessary, as the components of the KOOS and the scoring instructions are self-explanatory.

Translations/adaptations—Available in English and Swedish (original versions developed concurrently), Austria-German, Czech, Chinese, Croatian, Danish, Dutch, Estonian, French, German, Italian, Japanese, Latvian, Lithuanian, Norwegian, Persian, Portuguese, Polish, Russian, Singapore English, Slovak, Slovenian, Spanish (US), Spanish (Peru), Thai, Turkish, and Ukrainian. Cross-cultural adaptations have been conducted for the Swedish (27,28), Chinese (29), Dutch (30), French (31), Persian (32), Portuguese (33), Russian (Golubev; www.koos.nu), Singapore English (29), Thai (34), and Turkish (35) translations.

Psychometric Information

Method of development—Items were selected based on: 1) the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), version 3.0; 2) a literature review; 3) an expert panel (patients referred to physical therapy for knee injuries, orthopedic surgeons, and physical therapists from Sweden and the US); and 4) a pilot study of 2 questionnaires (1 for symptoms of ACL injury, 1 for symptoms of OA) in individuals with posttraumatic OA. Item-response theory was not used in the development of KOOS or for item selection (25).

Acceptability—Reported rates of missing data are low: 0.8% of items in patients who have undergone knee arthroscopy (27) and 3.2% of items on the pain, symptoms, ADL, and QOL subscales in patients prior to TKR (28). However, patients scheduled for TKR have also exhibited high rates of "not applicable" or missing items (74%) on the sport/recreation subscale (28). Studies consistently report no or acceptable floor or ceiling effects in knee injury cohorts (27,32,36) and in patients with mild or moderate knee OA (28,29,31,33). In those with severe OA awaiting TKR (28–31,33), 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%) sub-scales up to 12 months following TKR (28).

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. In knee OA, pain and ADL subscales have adequate test–retest data, while for the other subscales, reports indicate both lower and adequate test–retest reliability. Across the 5 subscales, the minimal detectable change ranges from 6–12 for knee injuries and from 13.4–21.1 for knee OA. 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 (25,28).

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 (25,27–30,32,33,36,37). 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 (38). A more recent study reported that the KOOS subscales had acceptable dimensionality (37).

Ability to detect change—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. Large effect sizes are consistently reported on all subscales 3–12 months after TKR. The minimum clinically important difference (MCID) and patient-acceptable symptom state (PASS) have not been calculated in any patient population.

Critical Appraisal of Overall Value to the Rheumatology Community

Strengths—The KOOS has undergone a substantial amount of psychometric testing, largely among populations for whom the scale was intended. 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 ACL, meniscal, and cartilage injury. 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 followup is necessary. Rasch analysis suggests that only the subscales that are not based on the WOMAC exhibit unidimensionality in patients who have undergone ACL reconstruction. 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 (28). The MCID and PASS are lacking from psychometric evaluation.

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 lack of reported MCID in any knee condition is a weakness.

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.

Intended populations/conditions: Knee osteoarthritis (OA).

Version: No modifications since the original publication (39).

Content—Measure of physical function derived from the activities of daily living and sport/recreation subscales of the KOOS (39). 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.

Endorsements—Osteoarthritis Research Society International and Outcome Measures in Rheumatology Clinical Trials.

Examples of use—Conditions: knee OA. Interventions: total knee replacement (TKR), intraarticular hyaluronic acid injection, and physical therapy.

Practical Application

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

Method of administration—Patient-completed questionnaire. Has not been validated for use during in-person or telephone interview.

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 (39).

Missing values: 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 (39). Training is not necessary, as the questionnaire and scoring instructions are self-explanatory.

Translations/adaptations—Available in English, Swedish, French, and Portuguese. 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 (40) and Portuguese (41) 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 (39).

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) (40).

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.

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 sub-scales; WOMAC function subscale (excluding KOOS-PS items); and Osteoarthritis Knee and Hip Quality of Life questionnaire (OAKHQOL) physical activity domain (40–42). 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 (40–42).

Ability to detect change—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 (41). The minimum clinically important difference (MCID) and patient-acceptable symptom state have not been reported.

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 has only undergone psychometric testing for this patient group. The MCID has not been reported.

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 (43).

Intended populations/conditions: Patients undergoing physical therapy for various knee pathologies, such as ligament/meniscal injury, osteoarthritis (OA), and patello-femoral pain (43–45). It is applicable for patients undergoing a variety of orthopedic knee procedures and young athletic subjects as well as older adults (46,47).

Version: Although originally described as a single index with 17 items (43), shorter versions have been widely used. A version using Likert-type scales is also available (48).

Content—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) (43,48). A separate scale has been developed to assess sporting activities (43).

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 (43,48).

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.

Endorsements—None.

Examples of use—Conditions: anterior cruciate ligament (ACL) injury, cartilage lesions, patellofemoral pain syndrome (PFPS), knee dislocation, and OA. Interventions: physical therapy, knee braces, ACL reconstruction, autologous chondrocyte implantation, patellar realignment surgery, and total knee replacement (TKR).

Practical Application

How to obtain—Presented in full as an appendix in the original publication (43).

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 (43,48).

Missing values: While there are no instructions provided as to handling missing data, the original publication only analyzed questionnaires with no missing data (43).

Score interpretation—Possible transformed score range 0–100, where 100 = no knee-related symptoms or functional limitations.

Normative values: Not available.

Respondent burden—It takes approximately 5 minutes to complete the KOS-ADL questionnaire (43). No training or assistance is required as the KOS-ADL is self-explanatory.

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

Translations/adaptations—The KOS-ADL instrument has been validated after translation to German (49), Portuguese (50), Turkish (51), and Greek (52).

Psychometric Information

Method of development—Initial item selection was conducted by review of existing patient-reported outcomes (e.g., Cincinnati Knee Scale, Lysholm Knee Scoring Scale, and Western Ontario and McMaster Universities Osteoarthritis Index [WOMAC]) and International Knee Documentation Committee guidelines. The list of items was modified by 12 physical therapists specialized in rehabilitation of musculoskeletal diseases of the knee (43).

Acceptability—No floor effects have been detected (46,47). Acceptable ceiling effects have been reported in people with a variety of knee pathologies undergoing physical therapy and orthopedic surgeon evaluation (43,47). However, high ceiling effects have been reported 6 months after TKR (46).

Reliability—In patients with mixed knee pathologies, the KOS-ADL has demonstrated adequate internal consistency across multiple languages, as well as adequate test–retest reliability for use in groups and individuals (Table 1).

Validity

Face and content validity: During development, the KOS-ADL 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 (43). 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 (43).

Construct validity: The KOS-ADL shows good correlation with other knee-specific scales, such as the Lysholm Knee Scoring Scale (43), WOMAC subscales (46), and global assessment of function (43). Higher correlations with the physical than mental component score of the Short Form 12 indicates convergent and divergent construct validity (46).

Ability to detect change—The KOS-ADL 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 (43). Moderate effect sizes were reported among patients with PFPS, with a minimum clinically important difference of 7.1 (45). Large effect sizes have been reported following TKR (46). The patient-acceptable symptom state has not been reported.

Critical Appraisal of Overall Value to the Rheumatology Community

Strengths—The KOS-ADL scale 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-ADL 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-ADL does not include items pertaining to athletic activities, such as running and jumping.

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

Research usability—The KOS-ADL 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

Description

Purpose—To evaluate outcomes of knee ligament surgery, particularly symptoms of instability (53).

Intended populations/conditions: Patients with knee ligament injury and anteromedial, anterolateral, combined anteromedial/anterolateral, posterolateral rotatory, or straight posterior instability (53).

Version: First published in 1982 (53). 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 (54).

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 (53). 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 (54).

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) (54).

Recall period for items—Not specified.

Endorsements—None.

Examples of use—Conditions: knee ligament injury (anterior cruciate ligament [ACL], posterior cruciate ligament [PCL], medial collateral ligament [MCL], lateral collateral ligament [LCL]), meniscal tears, knee cartilage lesions, osteochondritis dissecans, traumatic knee dislocation, patellar instability, patellofemoral pain, and knee osteoarthritis. Interventions: knee arthroscopy, ligament reconstruction (ACL, PCL, MCL, LCL), meniscal repair, meniscectomy, microfracture, osteochondral autografts, high tibial osteotomy, patellar realignment and stabilization surgery, lateral release, intraarticular hyaluronic acid injection, and therapeutic exercise.

Practical Application

How to obtain—The revised version is freely available in the publication (54). Multiple web sites publish versions of the scale, although they tend to differ slightly.

Method of administration—Original and revised scales were intended for in-person clinician administration (administered by the orthopedic surgeon with the patient's collaboration) (53,54), although subsequent studies have documented using the scale as a patient-completed questionnaire (55). While significantly lower scores have been found for questionnaires versus interview administration, suggesting interview bias (56), 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 (57).

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) (54).

Normative values: Normative data are available with and without stratification by sex (58,59).

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—Published in English. Although it has been used in international studies, no cross-cultural adaptations have been published.

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 (60). The authors added the item for instability, as they deemed this to be an important component of the disability associated with ACL injury (53). The revised scale does not report how the item for locking was selected (54). 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 (53). 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) (47,55,61–64).

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 (47). 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 (63–65). 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 (47,55). 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 (57).

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. The minimum clinically important difference (MCID) and patient-acceptable symptom state (PASS) have not been calculated in any patient population.

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.

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 ligament and meniscal injuries, chondral injuries, and patellar dislocation. It is important that researchers consistently utilize the same scale version (54). 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.

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 (66).

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 nonintuitive (i.e., lower scores represented better outcome, higher scores represented worse outcome), where the original 12 items are used but the scoring is different (67).

Content—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).

Number of items—12 items.

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

Recall period for items—Previous 4 weeks.

Endorsements—None.

Examples of use—Conditions: cartilage defects, tibio-femoral osteoarthritis (OA), patellofemoral OA, and rheumatoid arthritis. Interventions: autologous chondrocyte implantation, high tibial osteotomy, unicompartmental knee replacement, and TKR.

Practical Application

How to obtain—The original version can be found in its original publication (66). The modified version is freely available online (www.orthopaedicscore.com/scorepages/oxford_knee_score.html) (67).

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 (66), while in the modified version the total score ranges from 0–48 (67). 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 burden—Reported to involve minimal respondent burden (66). 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 (66). Calculation of the total score takes 1–5 minutes. No training is necessary.

Translations/adaptations—Translated and validated in many languages, including Chinese (68), German (69), Japanese (70), Swedish (71), and Thai (72).

Psychometric Information

Method of development—Item generation and reduction was conducted by interviewing patients considering TKR (66).

Acceptability—When tested in patients undergoing TKR, no missing data were reported preoperatively, while postoperative rates of missing data remained low (5%) (66). A more recent study reported no missing data before and 6 months after TKR (46). This study also reported no floor or ceiling effects prior to TKR. Six months postoperatively, although there were no floor effects, there were ceiling effects reported (27% of patients scored the top score).

Reliability—The OKS has adequate internal consistency across multiple languages (66,68–72) (Table 1). The original study reported adequate test–retest reliability for use in groups and individuals (66).

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 (66). Convergent and divergent construct validity is demonstrated by higher correlations with the Short Form 12 physical than mental component (46). The OKS has been shown to fit Rasch models following rescoring of some items (73), and removal of items for limp and kneeling (74).

Ability to detect change—The OKS demonstrates good sensitivity and responsiveness to change (Table 2). Large effect sizes have been reported 6–12 months after TKR (66,75). The OKS has also been found to be a good predictor of revision TKR within 6 months (76). The minimum clinically important difference (MCID) and patient-acceptable symptom state have not been reported.

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 (77).

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.

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. The lack of MCID is a weakness.

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) (78,79).

Intended populations/conditions: Patients with knee and hip OA (78,79).

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 (80,81). Reduced versions of the WOMAC have been validated but are not endorsed on the WOMAC web site (82–84).

Content—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” (78,79).

Recall period for items—48 hours.

Endorsements—Osteoarthritis Research Society International.

Examples of use—Conditions: knee OA, chondral defects, and anterior cruciate ligament (ACL) deficiency. Interventions: physical therapy, massage, self-management, group education, weight loss, exercise, hydrotherapy, Tai Chi, yoga, diet, knee braces, foot orthoses, electrotherapy (e.g., transcutaneous electrical nerve stimulation, laser, pulsed electrical stimulation), acupuncture, pharmacotherapy (drugs, supplements), corticosteroid injection, intraarticular hyaluronic acid injection, arthroscopy, autologous chondrocyte implantation, ACL reconstruction, and total knee replacement (TKR).

Practical Application

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.org.

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 (79,85–88).

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 (78,79).

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 (78).

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 (89).

Respondent burden—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 >80 languages (80), and has validated language translations for Arabic (90), Chinese (91), Finnish (92), German

(93), Hebrew (94), Italian (95), Japanese (96), Korean (97), Moroccan (98), Singapore (99), Spanish (100), Swedish (101,102), Thai (103), and Turkish (104,105).

Psychometric Information

Method of development—Items were generated by survey of patients with knee or hip OA, review of existing questionnaires (e.g., Health Assessment Questionnaire, Arthritis Impact Measurement Scales), and input from rheumatologists and epidemiologists with experience in clinical assessment of rheumatic diseases. Patients were also utilized in item reduction (78).

Acceptability—The original study and subsequent studies have reported low rates of missing data (46,78). Reports of floor and ceiling effects have differed between studies (46,91,103,105,106). The stiffness subscale has been reported as having floor and ceiling effects prior to intervention (46,91,105). Ceiling effects have been reported by various studies for all subscales 6 months and 2 years after TKR (46,106).

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. The minimal detectable change and standard error of the measure vary according to condition and subscale.

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 (91,94,95,98,104,105,107,108), while lower correlations with measures such as the SF-36 mental subscales indicate divergent construct validity (91,95,104,105,109). 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 (110). While 1 study found the pain subscale to demonstrate good item separation and unidimensionality in patients with knee or hip OA (111), 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 (110). 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 (111). Similarly, Davis et al (110) 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.

Ability to detect change—The WOMAC appears to be responsive to change following surgical and nonsurgical interventions for knee OA and chondral defects (Table 2). In patients with knee OA, large effect sizes are consistently reported on all 3 subscales up to 2 years post-TKR. 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 sub-scales. 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. 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). 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 (112).

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.

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.

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) (113).

Intended populations/conditions: Various knee conditions, including ligament, meniscus, and chondral injury; patellofemoral pain; osteochondritis dissecans; trabecular fracture; and iliotibial band syndrome (113).

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 (113).

Method of administration—Patient-completed questionnaire. It has not 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 (113).

Missing values: No specific instructions for handling missing values.

Score interpretation—The total possible score range is 0–16, where 16 = 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 (113).

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

Translations/adaptations—None.

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 (113).

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 (113) (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 (113).

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 (113).

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, this scale is most useful as an adjunct to other scales that assess other domains of knee function (114). 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 (113).

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.

TEGNER ACTIVITY SCORE (TAS)

Description

Purpose—To provide a standardized method of grading work and sporting activities (54). Developed to complement the Lysholm scale, based on observations that limitations in function scores (Lysholm) may be masked by a decrease in activity level (54).

Intended populations/conditions: Intended for use in conjunction with the Lysholm Knee Scoring Scale, originally in patients with anterior cruciate ligament (ACL) injury (54).

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 (54). Activity levels 6–10 can only be achieved if the person participates in recreational or competitive sport.

Recall period for items—Current ability.

Endorsements—None.

Examples of use—Conditions: knee ligament injury (ACL, posterior cruciate ligament [PCL], medial collateral ligament [MCL], lateral collateral ligament [LCL]), meniscal tears, knee cartilage lesions, osteochondritis dissecans, traumatic knee dislocation, patellar instability, patello-femoral pain, and knee osteoarthritis (OA). Interventions: knee arthroscopy, ligament reconstruction (ACL, PCL, MCL, LCL), meniscal repair, meniscectomy, microfracture, osteochondral autografts, high tibial osteotomy, patellar realignment and stabilization surgery, lateral release, intraarticular hyaluronic acid injection, and therapeutic exercise.

Practical Application

How to obtain—Freely available in the original publication (54).

Method of administration—Originally established as an in-person, clinician-administered tool (115), but has been used more recently as a patient-completed questionnaire (55,116).

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 (58).

Respondent burden—Reported to take mean ± SD 3.3 ± 0.6 minutes to complete in those who have undergone total knee replacement (117). 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 ($r = 0.25$, $P = 0.03$) (117).

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. Although it has been used in international studies, no cross-cultural adaptations have been published. 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.

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) (55,61,64,117).

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 (55,61,117). 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 (55,61,64,117).

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.

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.

Acknowledgments

Dr. Collins's work was supported by a National Health and Medical Research Council (Australia) Health Professional Research Training (Post-Doctoral) Fellowship. Dr. Roos's work was supported by the Swedish Medical Research Council.

References

1. Garratt AM, Brealey S, Gillespie WJ. Patient-assessed health instruments for the knee: a structured review. Rheumatology (Oxford). 2004; 43:1414–23. [PubMed: 15316121]
2. Veenhof C, Bijlsma JW, van den Ende CH, van Dijk GM, Pisters MF, Dekker J. Psychometric evaluation of osteoarthritis questionnaires: a systematic review of the literature. Arthritis Rheum. 2006; 55:480–92. [PubMed: 16739188]
3. Paxton EW, Fithian DC. Outcome instruments for patellofemoral arthroplasty. Clin Orthop Relat Res. 2005; 436:66–70. [PubMed: 15995422]
4. Streiner, DL.; Norman, GR. Health measurement scales: a practical guide to their development and use. 4. Oxford: Oxford University Press; 2008.
5. Roos EM, Engelhart L, Ranstam J, Anderson AF, Irrgang J, Marx RG, et al. ICRS recommendation document: patient-reported outcome instruments for use in patients with articular cartilage defects. Cartilage. 2011; 2:122–36.
6. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res. 1995; 4:293–307. [PubMed: 7550178]
7. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007; 60:34–42. [PubMed: 17161752]
8. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010; 63:737–45. [PubMed: 20494804]
9. Cohen, J. Statistical power analysis for the behavioral sciences. 2. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988.
10. Kvien TK, Heiberg T, Hagen KB. Minimal clinically important improvement/difference (MCII/MCID) and patient acceptable symptom state (PASS): what do these concepts mean? Ann Rheum Dis. 2007; 66(Suppl):iii40–1. [PubMed: 17934093]
11. Irrgang JJ, Anderson AF, Boland AL, Harner CD, Kurosaka M, Neyret P, et al. Development and validation of the International Knee Documentation Committee subjective knee form. Am J Sports Med. 2001; 29:600–13. [PubMed: 11573919]
12. Hefti F, Muller W, Jakob RP, Staubli HU. Evaluation of knee ligament injuries with the IKDC form. Knee Surg Sports Traumatol Arthrosc. 1993; 1:226–34. [PubMed: 8536037]

13. Anderson, AF. Rating scales. In: Fu, FH.; Harner, CD.; Vince, KL., editors. Knee surgery. Baltimore: Williams & Wilkins; 1994. p. 275-96.
14. 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. [PubMed: 16219941]
15. 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. [PubMed: 15483542]
16. Metsavaht L, Leporace G, Riberto M, de Mello Sposito MM, Batista LA. Translation and cross-cultural adaptation of the Brazilian version of the International Knee Documentation Committee subjective knee form: validity and reproducibility. Am J Sports Med. 2010; 38:1894–9. [PubMed: 20472755]
17. Fu SN, Chan YH. Translation and validation of Chinese version of International Knee Documentation Committee subjective knee form. Disabil Rehabil. 2011; 33:1186–9. [PubMed: 20969433]
18. Haverkamp D, Sierevelt IN, Breugem SJ, Lohuis K, Blankevoort L, van Dijk CN. Translation and validation of the Dutch version of the International Knee Documentation Committee subjective knee form. Am J Sports Med. 2006; 34:1680–4. [PubMed: 16816150]
19. Lertwanich P, Praphruetkit T, Keyurapan E, Lamsam C, Kulthanan T. Validity and reliability of Thai version of the International Knee Documentation Committee subjective knee form. J Med Assoc Thai. 2008; 91:1218–25. [PubMed: 18788694]
20. Crawford K, Briggs KK, Rodkey WG, Steadman JR. Reliability, validity, and responsiveness of the IKDC score for meniscus injuries of the knee. Arthroscopy. 2007; 23:839–44. [PubMed: 17681205]
21. Higgins LD, Taylor MK, Park D, Ghodadra N, Marchant M, Pietrobon R, et al. Reliability and validity of the International Knee Documentation Committee (IKDC) subjective knee form. Joint Bone Spine. 2007; 74:594–9. [PubMed: 17888709]
22. Agel J, LaPrade RF. Assessment of differences between the modified Cincinnati and International Knee Documentation Committee patient outcome scores: a prospective study. Am J Sports Med. 2009; 37:2151–7. [PubMed: 19633231]
23. Greco NJ, Anderson AF, Mann BJ, Cole BJ, Farr J, Nissen CW, et al. Responsiveness of the International Knee Documentation Committee subjective knee form in comparison to the Western Ontario and Mc-Master Universities Osteoarthritis Index, modified Cincinnati Knee Rating System, and Short Form 36 in patients with focal articular cartilage defects. Am J Sports Med. 2010; 38:891–902. [PubMed: 20044494]
24. Irrgang JJ, Anderson AF, Boland AL, Harner CD, Neyret P, Richmond JC, et al. Responsiveness of the International Knee Documentation Committee subjective knee form. Am J Sports Med. 2006; 34:1567–73. [PubMed: 16870824]
25. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS): development of a self-administered outcome measure. J Orthop Sports Phys Ther. 1998; 28:88–96. [PubMed: 9699158]
26. Paradowski PT, Bergman S, Sunden-Lundius A, Lohmander LS, Roos EM. Knee complaints vary with age and gender in the adult population: population-based reference data for the Knee injury and Osteoarthritis Outcome Score (KOOS). BMC Musculoskelet Disord. 2006; 7:38. [PubMed: 16670005]
27. Roos EM, Roos HP, Ekdahl C, Lohmander LS. Knee injury and Osteoarthritis Outcome Score (KOOS): validation of a Swedish version. Scand J Med Sci Sports. 1998; 8:439–48. [PubMed: 9863983]
28. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS): validation and comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes. 2003; 1:17. [PubMed: 12801417]
29. Xie F, Li SC, Roos EM, Fong KY, Lo NN, Yeo SJ, et al. Cross-cultural adaptation and validation of Singapore English and Chinese versions of the Knee injury and Osteoarthritis Outcome Score

(KOOS) in Asians with knee osteoarthritis in Singapore. Osteoarthritis Cartilage. 2006; 14:1098–103. [PubMed: 16814575]

30. De Groot IB, Favejee MM, Reijman M, Verhaar JA, Terwee CB. The Dutch version of the Knee Injury and Osteoarthritis Outcome Score: a validation study. Health Qual Life Outcomes. 2008; 6:16. [PubMed: 18302729]
31. Ornetti P, Parratte S, Gossec L, Tavernier C, Argenson JN, Roos EM, et al. Cross-cultural adaptation and validation of the French version of the Knee injury and Osteoarthritis Outcome Score (KOOS) in knee osteoarthritis patients. Osteoarthritis Cartilage. 2008; 16:423–8. [PubMed: 17905602]
32. Salavati M, Mazaheri M, Negahban H, Sohani SM, Ebrahimian MR, Ebrahimi I, et al. Validation of a Persian-version of Knee injury and Osteoarthritis Outcome Score (KOOS) in Iranians with knee injuries. Osteoarthritis Cartilage. 2008; 16:1178–82. [PubMed: 18411065]
33. Goncalves RS, Cabri J, Pinheiro JP, Ferreira PL. Cross-cultural adaptation and validation of the Portuguese version of the Knee injury and Osteoarthritis Outcome Score (KOOS). Osteoarthritis Cartilage. 2009; 17:1156–62. [PubMed: 19303082]
34. Chaipinyo K. Test-retest reliability and construct validity of Thai version of Knee Osteoarthritis Outcome Score (KOOS). Thai J Phys Ther. 2009; 31:67–76.
35. Paker N, Bugdayci D, Sabirli F, Ozel S, Ersoy S. Knee Injury and Osteoarthritis Outcome Score: reliability and validation of the Turkish version. Turkiye Klinikleri J Med Sci. 2007; 27:350–6.
36. Bekkers JE, de Windt TS, Raijmakers NJ, Dhert WJ, Saris DB. Validation of the Knee Injury and Osteoarthritis Outcome Score (KOOS) for the treatment of focal cartilage lesions. Osteoarthritis Cartilage. 2009; 17:1434–9. [PubMed: 19454278]
37. Salavati M, Akhbari B, Mohammadi F, Mazaheri M, Khorrami M. Knee injury and Osteoarthritis Outcome Score (KOOS): reliability and validity in competitive athletes after anterior cruciate ligament reconstruction. Osteoarthritis Cartilage. 2011; 19:406–10. [PubMed: 21255667]
38. Comins J, Brodersen J, Krogsgaard M, Beyer N. Rasch analysis of the Knee injury and Osteoarthritis Outcome Score (KOOS): a statistical re-evaluation. Scand J Med Sci Sports. 2008; 18:336–45. [PubMed: 18028282]
39. Perruccio AV, Stefan Lohmander L, Canizares M, Tennant A, Hawker GA, Conaghan PG, et al. The development of a short measure of physical function for knee OA KOOS-Physical Function Shortform (KOOS-PS): an OARSI/OMERACT initiative. Osteoarthritis Cartilage. 2008; 16:542–50. [PubMed: 18294869]
40. Ornetti P, Perruccio AV, Roos EM, Lohmander LS, Davis AM, Maillefert JF. Psychometric properties of the French translation of the reduced KOOS and HOOS (KOOS-PS and HOOS-PS). Osteoarthritis Cartilage. 2009; 17:1604–8. [PubMed: 19589400]
41. Goncalves RS, Cabri J, Pinheiro JP, Ferreira PL, Gil J. Reliability, validity and responsiveness of the Portuguese version of the Knee injury and Osteoarthritis Outcome Score-Physical Function Short-form (KOOS-PS). Osteoarthritis Cartilage. 2010; 18:372–6. [PubMed: 19912982]
42. Davis AM, Perruccio AV, Canizares M, Hawker GA, Roos EM, Maillefert JF, et al. Comparative, validity and responsiveness of the HOOS-PS and KOOS-PS to the WOMAC physical function subscale in total joint replacement for osteoarthritis. Osteoarthritis Cartilage. 2009; 17:843–7. [PubMed: 19215728]
43. Irrgang JJ, Snyder-Mackler L, Wainner RS, Fu FH, Harner CD. Development of a patient-reported measure of function of the knee. J Bone Joint Surg Am. 1998; 80:1132–45. [PubMed: 9730122]
44. Marx R. Knee rating scales. Arthroscopy. 2003; 19:1103–8. [PubMed: 14673453]
45. Piva SR, Gil AB, Moore CG, Fitzgerald GK. Responsiveness of the activities of daily living scale of the knee outcome survey and numeric pain rating scale in patients with patellofemoral pain. J Rehabil Med. 2009; 41:129–35. [PubMed: 19229444]
46. Impellizzeri F, Mannion A, Leunig M, Bizzini M, Naal F. Comparison of the reliability, responsiveness, and construct validity of 4 different questionnaires for evaluating outcomes after total knee arthroplasty. J Arthroplasty. 2010 E-pub ahead of print.
47. Marx RG, Jones EC, Allen AA, Altchek DW, O'Brien SJ, Rodeo SA, et al. Reliability, validity, and responsiveness of four knee outcome scales for athletic patients. J Bone Joint Surg Am. 2001; 83-A:1459–69. [PubMed: 11679594]

48. Irrgang, J. Development of a health related quality of life instrument to assess physical function related to pathology and impairment of the knee. Pittsburgh: University of Pittsburgh; 1999.
49. Bizzini M, Gorelick M. Development of a German version of the knee outcome survey for daily activities. Arch Orthop Trauma Surg. 2007; 127:781–9. [PubMed: 16865400]
50. Goncalves R, Cabri J, Pinheiro J. Cross-cultural adaptation and validation of the Portuguese version of the Knee Outcome Survey-Activities of Daily Living Scale (KOS-ADLS). Clin Rheumatol. 2008; 27:1445–9. [PubMed: 18773253]
51. Evcik D, Ay S, Ege A, Turel A, Kavuncu V. Adaptation and validation of Turkish version of the Knee Outcome Survey-Activities for Daily Living Scale. Clin Orthop Relat Res. 2009; 467:2077–82. [PubMed: 19350334]
52. Kapreli E, Panelli G, Strimpakos N, Billis E, Zacharopoulos A, Athanasopoulos S. Cross-cultural adaptation of the Greek version of the Knee Outcome Survey-Activities of Daily Living Scale (KOS-ADLS). Knee. 2010 E-pub ahead of print.
53. Lysholm J, Gillquist J. Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. Am J Sports Med. 1982; 10:150–4. [PubMed: 6896798]
54. Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. Clin Orthop Relat Res. 1985; 198:43–9. [PubMed: 4028566]
55. Briggs KK, Lysholm J, Tegner Y, Rodkey WG, Kocher MS, Steadman JR. The reliability, validity, and responsiveness of the Lysholm Score and Tegner Activity Scale for anterior cruciate ligament injuries of the knee: 25 years later. Am J Sports Med. 2009; 37:890–7. [PubMed: 19261899]
56. Hoher J, Bach T, Munster A, Bouillon B, Tiling T. Does the mode of data collection change results in a subjective knee score? Self-administration versus interview. Am J Sports Med. 1997; 25:642–7. [PubMed: 9302469]
57. Smith HJ, Richardson JB, Tennant A. Modification and validation of the Lysholm Knee Scale to assess articular cartilage damage. Osteoarthritis Cartilage. 2009; 17:53–8. [PubMed: 18556222]
58. Briggs KK, Steadman JR, Hay CJ, Hines SL. Lysholm score and Tegner activity level in individuals with normal knees. Am J Sports Med. 2009; 37:898–901. [PubMed: 19307332]
59. Demirdjian AM, Petrie SG, Guanche CA, Thomas KA. The outcomes of two knee scoring questionnaires in a normal population. Am J Sports Med. 1998; 26:46–51. [PubMed: 9474400]
60. Oretorp N, Gillquist J, Liljedahl SO. Long term results of surgery for non-acute anteromedial rotatory instability of the knee. Acta Orthop Scand. 1979; 50:329–36. [PubMed: 474103]
61. Briggs KK, Kocher MS, Rodkey WG, Steadman JR. Reliability, validity and responsiveness of the Lysholm Knee Score and the Tegner Activity Scale for patients with meniscal injury of the knee. J Bone Joint Surg Am. 2006; 88:698–705. [PubMed: 16595458]
62. Heintjes EM, Bierma-Zeinstra SM, Berger MY, Koes BW. Lysholm scale and WOMAC index were responsive in prospective cohort of young general practice patients. J Clin Epidemiol. 2008; 61:481–8. [PubMed: 18394542]
63. Kocher MS, Steadman JR, Briggs KK, Sterett WI, Hawkins RJ. Reliability, validity, and responsiveness of the Lysholm knee scale for various chondral disorders of the knee. J Bone Joint Surg Am. 2004; 86-A:1139–45. [PubMed: 15173285]
64. Paxton EW, Fithian DC, Stone ML, Silva P. The reliability and validity of knee-specific and general health instruments in assessing acute patellar dislocation outcomes. Am J Sports Med. 2003; 31:487–92. [PubMed: 12860533]
65. Sgaglione NA, Del Pizzo W, Fox JM, Friedman MJ. Critical analysis of knee ligament rating systems. Am J Sports Med. 1995; 23:660–7. [PubMed: 8600730]
66. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. J Bone Joint Surg Br. 1998; 80:63–9. [PubMed: 9460955]
67. Murray D, Fitzpatrick R, Rogers K, Pandit H, Beard D, Carr A, et al. The use of the Oxford hip and knee scores. J Bone Joint Surg Br. 2007; 89:1010–4. [PubMed: 17785736]
68. Xie F, Li S, Lo N, Yeo S, Yang K, Yeo W, et al. Cross-cultural adaptation and validation of Singapore English and Chinese Versions of the Oxford Knee Score (OKS) in knee osteoarthritis patients undergoing total knee replacement. Osteoarthritis Cartilage. 2007; 15:1019–24. [PubMed: 17408984]

69. Naal F, Impellizzeri F, Sieverding M, Loibl M, Von Knoch F, Mannion A, et al. The 12-item Oxford Knee Score: cross-cultural adaptation into German and assessment of its psychometric properties in patients with osteoarthritis of the knee. Osteoarthritis Cartilage. 2009; 17:49–52. [PubMed: 18602843]
70. Takeuchi R, Sawaguchi T, Nakamura N, Ishikawa H, Saito T, Goldhahn S. Cross-cultural adaptation and validation of the Oxford 12-item knee score in Japanese. Arch Orthop Trauma Surg. 2011; 131:247–54. [PubMed: 20830479]
71. Dunbar M, Robertsson O, Ryd L, Lidgren L. Translation and validation of the Oxford-12 item knee score for use in Sweden. Acta Orthop Scand. 2000; 71:268–74. [PubMed: 10919298]
72. Charoencholvanich K, Pongcharoen B. Oxford knee score and SF-36: translation & reliability for use with total knee arthroscopy patients in Thailand. J Med Assoc Thai. 2005; 88:1194–202. [PubMed: 16536104]
73. Conaghan PG, Emerton M, Tennant A. Internal construct validity of the Oxford Knee Scale: evidence from Rasch measurement. Arthritis Rheum. 2007; 57:1363–7. [PubMed: 18050174]
74. Ko Y, Lo N, Yeo S, Yang K, Yeo W, Chong H, et al. Rasch analysis of the Oxford Knee Score. Osteoarthritis Cartilage. 2009; 17:1163–9. [PubMed: 19409293]
75. Garratt A, Brealey S, Gillespie W. Patient-assessed health instruments for the knee: a structured review. Rheumatology (Oxford). 2004; 43:1414–23. [PubMed: 15316121]
76. Moonot P, Medalla G, Matthews D, Kalairajah Y, Field R. Correlation between the Oxford Knee and American Knee Society scores at mid-term follow-up. J Knee Surg. 2009; 22:226–30. [PubMed: 19634726]
77. Medalla GA, Moonot P, Peel T, Kalairajah Y, Field RE. Cost-benefit comparison of the Oxford Knee Score and the American Knee Society Score in measuring outcome of total knee arthroplasty. J Arthroplasty. 2009; 24:652–6. [PubMed: 18617359]
78. Bellamy, N. WOMAC Osteoarthritis Index user guide. London (Ontario, Canada): University of Western Ontario; 1995.
79. Bellamy, N. WOMAC Osteoarthritis Index user guide. Brisbane (Australia): CONROD, The University of Queensland; 2002. Version V
80. WOMAC-AUSCAN-osteoarthritis global index. http://www.womac.org
81. Ornetti P, Dougados M, Paternotte S, Logeart I, Gossec L. Validation of a numerical rating scale to assess functional impairment in hip and knee osteoarthritis: comparison with the WOMAC function scale. Ann Rheum Dis. 2011; 70:740–6. [PubMed: 21149497]
82. Baron G, Tubach F, Ravaud P, Logeart I, Dougados M. Validation of a short form of the Western Ontario and McMaster Universities Osteoarthritis Index function subscale in hip and knee osteoarthritis. Arthritis Rheum. 2007; 57:633–8. [PubMed: 17471534]
83. Whitehouse SL, Lingard EA, Katz JN, Learmonth ID. Development and testing of a reduced WOMAC function scale. J Bone Joint Surg Br. 2003; 85:706–11. [PubMed: 12892194]
84. Yang KG, Raijmakers NJ, Verbout AJ, Dhert WJ, Saris DB. Validation of the short-form WOMAC function scale for the evaluation of osteoarthritis of the knee. J Bone Joint Surg Br. 2007; 89:50–6. [PubMed: 17259416]
85. Bellamy N, Campbell J, Stevens J, Pilch L, Stewart C, Mahmood Z. Validation study of a computerized version of the Western Ontario and McMaster Universities VA3. 0 Osteoarthritis Index. J Rheumatol. 1997; 24:2413–5. [PubMed: 9415651]
86. Bellamy N, Campbell J, Hill J. A comparative study of telephone vs on-site completion of the WOMAC 3. 0 Osteoarthritis Index. J Rheumatol. 2002; 29:783–6. [PubMed: 11950022]
87. Bellamy N, Wilson C, Hendrikz J, Whitehouse SL, Patel B, Dennison S, et al. Osteoarthritis Index delivered by mobile phone (m-WOMAC) is valid, reliable, and responsive. J Clin Epidemiol. 2011; 64:182–90. [PubMed: 20609562]
88. Theiler R, Speilberger J, Bischoff H, Bellamy N, Huber J, Kroesen S. Clinical evaluation of the WOMAC 3. 0 OA Index in numeric rating scale format using a computerised touch screen version. Osteoarthritis Cartilage. 2002; 10:479–81. [PubMed: 12056851]
89. Bellamy N, Wilson C, Hendrikz J. Population-based normative values for the Western Ontario and McMaster (WOMAC) Osteoarthritis Index and the Australian/Canadian (AUSCAN) hand

osteoarthritis index functional subscales. Inflammopharmacology. 2010; 18:1–8. [PubMed: 20024627]

90. Guermazi M, Poiraudeau S, Yahia M, Mezganni M, Fermanian J, Habib Elleuch M, et al. Translation, adaptation and validation of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) for an Arab population: the Sfax modified WOMAC. Osteoarthritis Cartilage. 2004; 12:459–68. [PubMed: 15135142]
91. Xie F, Li SC, Goeree R, Tarride JE, O'Reilly D, Lo NN, et al. Validation of Chinese Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) in patients scheduled for total knee replacement. Qual Life Res. 2008; 17:595–601. [PubMed: 18415706]
92. Soininen JV, Paavolainen PO, Gronblad MA, Kaapa EH. Validation study of a Finnish version of the Western Ontario and McMasters University Osteoarthritis Index. Hip Int. 2008; 18:108–11. [PubMed: 18645984]
93. Stucki G, Meier D, Stucki S, Michel BA, Tyndall AG, Dick W, et al. Evaluation of a German version of WOMAC (Western Ontario and McMaster Universities) Arthrosis Index. Z Rheumatol. 1996; 55:40–9. In German. [PubMed: 8868149]
94. Wigler I, Neumann L, Yaron M. Validation study of a Hebrew version of WOMAC in patients with osteoarthritis of the knee. Clin Rheumatol. 1999; 18:402–5. [PubMed: 10524555]
95. Salaffi F, Leardini G, Canesi B, Mannoni A, Fioravanti A, Caporali R, et al. Reliability and validity of the Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index in Italian patients with osteoarthritis of the knee. Osteoarthritis Cartilage. 2003; 11:551–60. [PubMed: 12880577]
96. Hashimoto H, Hanyu T, Sledge CB, Lingard EA. Validation of a Japanese patient-derived outcome scale for assessing total knee arthroplasty: comparison with Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). J Orthop Sci. 2003; 8:288–93. [PubMed: 12768467]
97. Bae SC, Lee HS, Yun HR, Kim TH, Yoo DH, Kim SY. Cross-cultural adaptation and validation of Korean Western Ontario and McMaster Universities (WOMAC) and Lequesne osteoarthritis indices for clinical research. Osteoarthritis Cartilage. 2001; 9:746–50. [PubMed: 11795994]
98. Faik A, Benbouazza K, Amine B, Maaroufi H, Bahiri R, Lazrak N, et al. Translation and validation of Moroccan Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index in knee osteoarthritis. Rheumatol Int. 2008; 28:677–83. [PubMed: 18092169]
99. Thumboo J, Chew LH, Soh CH. Validation of the Western Ontario and McMaster University Osteoarthritis Index in Asians with osteoarthritis in Singapore. Osteoarthritis Cartilage. 2001; 9:440–6. [PubMed: 11467892]
100. Escobar A, Quintana JM, Bilbao A, Azkarate J, Guenaga JI. Validation of the Spanish version of the WOMAC questionnaire for patients with hip or knee osteoarthritis: Western Ontario and McMaster Universities Osteoarthritis Index. Clin Rheumatol. 2002; 21:466–71. [PubMed: 12447629]
101. Roos EM, Klassbo M, Lohmander LS. WOMAC osteoarthritis index: reliability, validity, and responsiveness in patients with arthroscopically assessed osteoarthritis. Western Ontario and McMaster Universities. Scand J Rheumatol. 1999; 28:210–5. [PubMed: 10503556]
102. Soderman P, Malchau H. Validity and reliability of Swedish WOMAC Osteoarthritis Index: a self-administered disease-specific questionnaire (WOMAC) versus generic instruments (SF-36 and NHP). Acta Orthop Scand. 2000; 71:39–46. [PubMed: 10743991]
103. Kuptniratsaikul V, Rattanachaiyanont M. Validation of a modified Thai version of the Western Ontario and McMaster (WOMAC) Osteoarthritis Index for knee osteoarthritis. Clin Rheumatol. 2007; 26:1641–5. [PubMed: 17310271]
104. Basaran S, Guzel R, Seydaoglu G, Guler-Uysal F. Validity, reliability, and comparison of the WOMAC osteoarthritis index and Lequesne algofunctional index in Turkish patients with hip or knee osteoarthritis. Clin Rheumatol. 2010; 29:749–56. [PubMed: 20169459]
105. Tuzun EH, Eker L, Aytar A, Daskapan A, Bayramoglu M. Acceptability, reliability, validity and responsiveness of the Turkish version of WOMAC osteoarthritis index. Osteoarthritis Cartilage. 2005; 13:28–33. [PubMed: 15639634]

106. Escobar A, Quintana JM, Bilbao A, Arostegui I, Lafuente I, Vidaurreta I. Responsiveness and clinically important differences for the WOMAC and SF-36 after total knee replacement. Osteoarthritis Cartilage. 2007; 15:273–80. [PubMed: 17052924]

107. Brazier J, Harper R, Munro J, Walters S, Snaith M. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. Rheumatology (Oxford). 1999; 38:870–7. [PubMed: 10515649]

108. McConnell S, Kolopack P, Davis AM. The Western Ontario and Mc-Master Universities Osteoarthritis Index (WOMAC): a review of its utility and measurement properties. Arthritis Rheum. 2001; 45:453–61. [PubMed: 11642645]

109. Bombardier C, Melfi C, Paul J, Hawker G, Wright J, Coyte P. Comparison of a generic and a disease-specific measure of pain and physical function after knee replacement surgery. Med Care. 1995; 33(Suppl):AS131–44. [PubMed: 7723441]

110. Davis A, Badley E, Beaton D, Kopec J, Wright J, Young N, et al. Rasch analysis of the Western Ontario McMaster (WOMAC) Osteoarthritis Index: results from community and arthroplasty samples. J Clin Epidemiol. 2003; 56:1076–83. [PubMed: 14614998]

111. Wolfe F, Kong SX. Rasch analysis of the Western Ontario McMaster questionnaire (WOMAC) in 2,205 patients with osteoarthritis, rheumatoid arthritis, and fibromyalgia. Ann Rheum Dis. 1999; 58:563–8. [PubMed: 10460190]

112. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. Ann Rheum Dis. 2005; 64:34–7. [PubMed: 15130902]

113. Marx R, Stump T, Jones E, Wickiewicz T, Warren R. Development and evaluation of an activity rating scale for disorders of the knee. Am J Sports Med. 2001; 29:213–8. [PubMed: 11292048]

114. Rick W. Knee injury outcomes measures. J Am Acad Orthop Surg. 2009; 17:31–9. [PubMed: 19136425]

115. Hambly K. The use of the Tegner Activity Scale for articular cartilage repair of the knee: a systematic review. Knee Surg Sports Traumatol Arthrosc. 2011; 19:604–14. [PubMed: 21076815]

116. Frobell RB, Roos EM, Roos HP, Ranstam J, Lohmander LS. A randomized trial of treatment for acute anterior cruciate ligament tears. N Engl J Med. 2010; 363:331–42. [PubMed: 20660401]

117. Naal FD, Impellizzeri FM, Leunig M. Which is the best activity rating scale for patients undergoing total joint arthroplasty? Clin Orthop Relat Res. 2009; 467:958–65. [PubMed: 18587624]

118. Frobell RB, Svensson E, Gothrick M, Roos EM. Self-reported activity level and knee function in amateur football players: the influence of age, gender, history of knee injury and level of competition. Knee Surg Sports Traumatol Arthrosc. 2008; 16:713–9. [PubMed: 18350275]

119. Bengtsson J, Mollborg J, Werner S. A study for testing the sensitivity and reliability of the Lysholm knee scoring scale. Knee Surg Sports Traumatol Arthrosc. 1996; 4:27–31. [PubMed: 8819060]

120. Marx RG, Menezes A, Horovitz L, Jones EC, Warren RF. A comparison of two time intervals for test-retest reliability of health status instruments. J Clin Epidemiol. 2003; 56:730–5. [PubMed: 12954464]

121. Dunbar MJ, Robertsson O, Ryd L, Lidgren L. Appropriate questionnaires for knee arthroplasty: results of a survey of 3,600 patients from The Swedish Knee Arthroplasty Registry. J Bone Joint Surg Br. 2001; 83:339–44. [PubMed: 11341416]

122. Fransen M, Edmonds J. Reliability and validity of the EuroQol in patients with osteoarthritis of the knee. Rheumatology (Oxford). 1999; 38:807–13. [PubMed: 10515639]

123. Stucki G, Sangha O, Stucki S, Michel BA, Tyndall A, Dick W, et al. Comparison of the WOMAC (Western Ontario and McMaster Universities) Osteoarthritis Index and a self-report format of the self-administered Lequesne-Algofunctional index in patients with knee and hip osteoarthritis. Osteoarthritis Cartilage. 1998; 6:79–86. [PubMed: 9692062]

124. Angst F, Ewert T, Lehmann S, Aeschlimann A, Stucki G. The factor subdimensions of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) help to specify hip

and knee osteoarthritis: a prospective evaluation and validation study. J Rheumatol. 2005; 32:1324–30. [PubMed: 15996072]

125. Chesworth BM, Mahomed NN, Bourne RB, Davis AM. Willingness to go through surgery again validated the WOMAC clinically important difference from THR/TKR surgery. J Clin Epidemiol. 2008; 61:907–18. [PubMed: 18687289]
126. Davis AM, Lohmander LS, Wong R, Venkataramanan V, Hawker GA. Evaluating the responsiveness of the ICOAP following hip or knee replacement. Osteoarthritis Cartilage. 2010; 18:1043–5. [PubMed: 20435154]
127. Theiler R, Bischoff-Ferrari HA, Good M, Bellamy N. Responsiveness of the electronic touch screen WOMAC 3. 1 OA Index in a short term clinical trial with rofecoxib. Osteoarthritis Cartilage. 2004; 12:912–6. [PubMed: 15501407]
128. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. Ann Rheum Dis. 2005; 64:29–33. [PubMed: 15208174]

Table 1

Summary of reliability data*

	Patient cohort evaluated (ref.)	Internal consistency (Cronbach's *α*)	Test–retest (ICC)	MDC	SEM
Function measures					
IKDC	Knee injuries (ACL, meniscal, chondral) (15,20,23)	0.77–0.91	0.90–0.95†	8.8–15.6†	3.2–5.6†
	Cohort of mixed knee pathologies (11,16–18,21)	0.92–0.97	0.87–0.99†	6.7	2.4–4.6†
KOOS	Knee injuries (25,27,32,36)	Pain: 0.84–0.91 Symptoms: 0.25–0.75 ADL: 0.94–0.96 Sport/rec: 0.85–0.89 QOL: 0.64–0.9	Pain: 0.85–0.93 Symptoms: 0.83–0.95 ADL: 0.75–0.91 Sport/rec: 0.61–0.89 QOL: 0.83–0.95	Pain: 6–6.1 Symptoms: 5–8.5 ADL: 7–8 Sport/rec: 5.8–12 QOL: 7–7.2	Pain: 2.2 Symptoms: 3.1 ADL: 2.9 Sport/rec: 2.1 QOL: 2.6
	Knee OA (28–31,33)	Pain: 0.65–0.94 Symptoms: 0.56–0.83 ADL: 0.78–0.97 Sport/rec: 0.84–0.98 QOL: 0.71–0.85	Pain: 0.8–0.97 Symptoms: 0.74–0.94 ADL: 0.84–0.94 Sport/rec: 0.65–0.92 QOL: 0.6–0.91	Pain: 13.4 Symptoms: 15.5 ADL: 15.4 Sport/rec: 19.6 QOL: 21.1	Pain: 7.2–10.1 Symptoms: 7.2–9 ADL: 5.2–11.7 Sport/rec: 9–24.6 QOL: 7.4–10.8
KOOS-PS	Knee OA (40–42)	0.89	0.85–0.86	–	–
KOS-ADL	Mixed knee pathologies (43,47,49–52)	0.89–0.98	0.94–0.98	11.4	4.1
Lysholm Knee Scoring Scale	Knee injuries (ACL, meniscal, chondral; patellar dislocation) (54,55,61,63,64)	0.65–0.73	0.88–0.97	8.9–10.1	3.2–3.6
	Mixed knee pathologies (43,47,119,120)	0.60–0.73	0.68–0.95	–	9.7–12.5†
OKS	Knee OA (46,66,71,121)	0.87–0.93	0.91–0.94	6.1	2.2
WOMAC	Chondral defects (23)		Pain: 0.81–0.85 Symptoms: 0.75–0.86 Function: 0.86–0.93	Pain: 14.4–16.2 Symptoms: 22.9–30.6 Function: 10.6–15	Pain: 5.2–5.8 Symptoms: 8.3–11.1 Function: 3.8–5.4
	Knee OA (42,46,91,92, 94–98,100,101,103–105,108,122,123)	Pain: 0.67–0.92 Symptoms: 0.7–0.94 Function: 0.82–0.98	Pain: 0.65–0.98 Symptoms: 0.52–0.89 Function: 0.71–0.96	Pain: 18.8–22.4 Symptoms: 27.1–29.1 Function: 13.1–13.3	Pain: 6.8–8.1 Symptoms: 9.8–10.5 Function: 4.7–4.8
Activity measures					
ARS	Baseline knee athletic activity for cohort of mixed knee pathologies (113)	–	0.97	–	–
TAS	Knee injuries (ACL, meniscal patellar dislocation) (55,61,64)	n/a	0.82–0.92†	1.0	0.4–0.64
	Knee OA (117)	n/a	0.84	–	–

*ICC = intraclass correlation coefficient; MDC = minimal detectable change; SEM = standard error of measurement; IKDC = International Knee Documentation Committee Subjective Knee Evaluation Form; ACL = anterior cruciate ligament; KOOS = Knee Injury and Osteoarthritis Outcome Score; ADL = activities of daily living; sport/rec = sport/recreation; QOL = quality of life; OA = osteoarthritis; KOOS-PS = Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form; KOS-ADL = Knee Outcome Survey Activities of Daily Living Scale; OKS = Oxford Knee Score; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index; ARS = Activity Rating Scale; TAS = Tegner Activity Scale; n/a = not applicable.

† Large variation in time between test—retest (up to 12 months).

Table 2

Summary of responsiveness data*

	Patient cohort evaluated	ES	SRM	MCID
Function measures				
IKDC	Knee injuries (ACL, meniscal, chondral) (20,23)	Meniscal repair/resection (12 m): 2.11 Various cartilage procedures: 0.76 (6 m), 1.06 (12 m)	Meniscal repair/resection (12 m): 1.5 Various cartilage procedures: 0.57 (6 m), 1.0 (12 m)	Chondral injuries: 6.3 (6 m), 16.7 (12 m)
	Cohort of mixed knee pathologies (22,24)	Various surgical procedures (6–28 m): 1.13	Various surgical procedures: 4.4 (4–8 m), 0.94 (6–28 m)	6–28 m: 11.5 (sensitive), 20.5 (specific)
KOOS	Knee injuries (25,27,36)	Partial meniscectomy (3 m): 1.11 (pain), 0.93 (symp.), 0.67 (ADL), 0.9 (sport/rec), 1.15 (QOL) ACLR (6 m): 0.84 (pain), 0.87 (symp.), 0.94 (ADL), 1.16 (sport/rec), 1.65 (QOL) ACI, MF (3 y): 0.82 (pain), 0.72 (symp.), 0.7 (ADL), 0.98 (sport/rec), 1.32 (QOL)	ACI, MF (3 y): 0.71 (pain), 0.61 (symp.), 0.75 (ADL), 0.87 (sport/rec), 0.76 (QOL)	–
	Knee OA (28,31,33)	PT (4 w): 1.08 (pain), 0.97 (symp.), 1.07 (ADL), 0.79 (sport/rec), 0.78 (QOL) TKR (3 m): 2.59 (pain), 1.63 (symp.), 2.52 (ADL), 1.31 (sport/rec), 2.8 (QOL) TKR (6 m): 2.28 (pain), 1.24 (symp.), 2.25 (ADL), 1.18 (sport/rec), 2.86 (QOL) TKR (12 m): 2.55 (pain), 1.59 (symp.), 2.56 (ADL), 1.08 (sport/rec), 3.54 (QOL)	PT (4 w): 1.28 (pain), 1.02 (symp.), 1.37 (ADL), 0.83 (sport/rec), 0.87 (QOL) TKR (3 m): 1.85 (pain), 1.45 (symp.), 1.8 (ADL), 0.89 (sport/rec), 1.93 (QOL) TKR (6 m): 1.67 (pain), 0.99 (symp.), 1.7 (ADL), 0.81 (sport/rec), 1.6 (QOL) TKR (12 m): 2.12 (pain), 1.25 (symp.), 1.9 (ADL), 0.88 (sport/rec), 1.99 (QOL)	–
KOOS-PS	Knee OA (40–42)	PT (4 w): 0.5–0.88 HAI (4 w): 0.51	PT (4 w): 0.73–1.21 HAI (4 w): 0.8 TKR (6 m): 1.4	–
KOS-ADL	Mixed knee pathologies (43,45–47)	PT: 0.44 (1 w), 0.94 (4 w), 1.26 (8 w) PT (6 w): 0.63 TKR (6 m): 1.3	PT (6 w): 7.1 TKR (6 m): 1.1	PFPS: 7.1
Lysholm Knee Scoring Scale	Knee injuries (ACL, meniscal, chondral; patellar dislocation) (55,61,63)	ACLR: 1.0 (6–9 m), 1.1 (1–2 y) Meniscal repair (1 y): 1.2 MF (1–6 y): 1.2	ACLR: 0.93 (6 m), 1.1 (9 m), 1.2 (1 y), 0.93 (2 y) Meniscal repair (1 y): 0.97–1.13 MF (1–6 y): 1.1	–
	Mixed knee pathologies (47,62,120)	PT (1 m): 0.9	Variety of nonsurgical and surgical interventions (3 m): 0.9	–
OKS	Knee OA (46,66)	TKR (6 m): 0.9–2.19	TKR (6 m): 0.7	–
WOMAC	Chondral defects (23)	Various cartilage surgeries (6 m): 0.98 (pain), 0.51 (symp.), 0.88 (function) Various cartilage surgeries (12 m): 1.14 (pain), 0.72 (symp.), 1.2 (function)	Various cartilage surgeries (6 m): 0.91 (pain), 0.40 (symp.), 0.86 (function) Various cartilage surgeries (12 m): 0.94 (pain), 0.64 (symp.), 1.13 (function)	–
	Knee OA (42,46,92, 96,97,100,101,105, 106,108,124–128)	TKR (3 m): 1.62 (pain), 1.26 (symp.), 2.02 (function) TKR (6 m): 0.95–1.9 (pain), 0.88–1.5 (symp.), 1.01–2.2 (function) TKR (1 y): 1.8–2.4 (pain), 1.8–3.1 (function) TKR (2 y): 1.9–41 (pain), 1.3–24 (symp.), 1.7–23.9 (function) Exercise (2 w): 0.74–0.88 (pain), 0.32–0.44 (symp.), 0.50–0.79 (function) Exercise (6 m): 0.41 (pain), 0.28 (function) Rehabilitation (not defined): 0.52 (pain), 0.42 (symp.), 0.44 (function) Drug (2 w): 0.94 (pain), 0.46 (symp.), 0.72 (function) Drug (3 w): 0.76–0.88 (pain), 0.59–0.63 (symp.), 0.75–0.77 (function) Drug (4 w): 0.69 (pain), 0.41 (symp.), 0.56 (function)	TKR (3 m): 1.14–1.58 (pain), 1.15 (symp.), 1.02–2.02 (function) TKR (6 m): 0.95–1.8 (pain), 0.63–1.3 (symp.), 0.9–1.9 (function) TKR (2 y): 1.55 (pain), 1.03 (symp.), 1.32 (function) Drug (2 w): 1.09 (pain), 0.43 (symp.), 0.89 (function) Exercise (2 w): 0.78–1 (pain), 0.29–0.52 (symp.), 0.69–0.94 (function)	NSAIDs (4 w, function): 9.1 (absolute), 26 (relative) TKR (6 m): 22.87 (pain), 14.43 (symp.), 19.01 (function) TKR (12 m): 36 (pain), 33 (function)

Arthritis Care Res (Hoboken). Author manuscript; available in PMC 2015 February 21.

	Patient cohort evaluated	ES	SRM	MCID
		Drug (6 w): 0.53–0.8 (pain), 0.6–0.75 (symp.), 0.58–0.82 (function) Drug (8 w): 0.58 (pain), 0.53 (symp.), 0.76 (function) Drug (12 w): 0.44–0.91 (pain), 0.55–0.84 (symp.), 0.58–0.81 (function) Acupuncture (3 w): 0.4 (pain), 0.52 (symp.), 0.31 (function) Acupuncture (8 w): 1.3 (pain), 1.2 (function)		TKR (2 y): 27.98 (pain), 21.35 (symp.), 20.84 (function)
Activity measures				
ARS	Baseline knee athletic activity for cohort of mixed knee pathologies	–	–	–
TAS	Knee injuries (ACL, meniscal; patellar dislocation) (55,61)	Various meniscal surgeries (12 m): 0.61 (isolated lesions), 0.84 (combined lesions) ACLR: 0.74 (6 m), 1.1 (9 m), 1.0 (1 y), 1.0 (2 y)	Various meniscal surgeries (12 m): 0.6 (isolated lesions), 0.7 (combined lesions) ACLR: 0.61 (6 m), 0.84 (9 m), 0.96 (1 y), 1.0 (2 y)	–
	Knee OA	–	–	–

*ES = effect size; SRM = standardized response mean; MCID = minimum clinically important difference; IKDC = International Knee Documentation Committee Subjective Knee Evaluation Form; ACL = anterior cruciate ligament; KOOS = Knee Injury and Osteoarthritis Outcome Score; symp. = symptoms; ADL = activities of daily living; sport/rec = sport/recreation; QOL = quality of life; ACLR = ACL reconstruction; ACI = autologous chondrocyte implantation; MF = microfracture; OA = osteoarthritis; PT = physical therapy; TKR = total knee replacement; KOOS-PS = Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form; HAI = intraarticular hyaluronic acid injection; KOS-ADL = Knee Outcome Survey Activities of Daily Living Scale; PFPS = patellofemoral pain syndrome; OKS = Oxford Knee Score; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index; NSAIDs = nonsteroidal antiinflammatory drugs; ARS = Activity Rating Scale; TAS = Tegner Activity Scale.

Summary Table for Knee Function Measures*

Scale	Purpose/content	Method of administration	Respondent burden	Administrative burden	Score interpretation	Reliability evidence	Validity evidence	Ability to detect change	Strengths	Cautions
Knee function										
IKDC	Symptoms, sport/daily activities, function; variety of knee conditions	Patient completed	10 min	5 min; manual scoring using guidelines provided	Single score; 0–100 (100 = no symptoms, no limitation with daily/sport activities)	Internal: adequate; test–retest: adequate for groups/individuals with knee injuries	Face: adequate; content: cannot be assumed; construct: adequate	Responsive to change following surgery; MCID for cartilage repair, various knee surgeries	No floor/ceiling effects	No patient input in development; long recall period; missing data; lacking psychometric testing in knee OA; aggregate score may mask deficits in 1 domain; multiple versions available
KOOS	Pain, symptoms, ADL, sport/rec, QOL; posttraumatic knee OA and preceding conditions	Patient completed	10 min	5 min; scoring spreadsheet	5 subscales; 0–100 (100 = no problems)	Internal, test–retest: variable (subscale, condition)	Face: adequate; content: adequate; construct: adequate	Responsive to change across a variety of knee conditions following surgical and nonsurgical interventions; MCID: NR	Substantial psychometric testing and cross-cultural validation; individual rather than aggregate scores	Not validated for interview administration; applicability of sport/rec items in older/less physically active patients
KOOS-PS	Function (ADL, sport/rec); knee OA	Patient completed	2 min	<5 min; conversion table	Single score; 0–100 (100 = no difficulty)	Internal: adequate; test–retest: adequate for groups; less than adequate for individuals	Face: adequate; content: adequate; construct: adequate	Responsive to change following physical therapy and hyaluronic acid injection; MCID: NR	Developed using Rasch analysis; minimal burden	Psychometric testing only in knee OA
KOS-ADL	Symptoms, functional limitations; various knee pathologies (ligament/meniscal injuries, OA, PFP)	Patient completed	5 min	<5 min; manual calculation	Single score; 0–100 (100 = no knee- related symptoms or functional limitations)	Internal: adequate; test–retest: adequate	Face: adequate; content: cannot be assumed; construct: adequate	Responsive to change across a variety of knee disorders and interventions (physical therapy, TKR); MCID for PFP	Reliable and valid	No patient input in development; descriptive responses may be confusing; ensure use of consistent version; may not be appropriate for highly active patients
Lysholm Knee Scoring Scale	Limp, support, locking, instability, pain, swelling, stairs, squatting; knee ligament surgery	In-person clinician administration	Variable depending on administration method	<5 min; manual calculation	Single score; 0–100 (100 = no symptoms or disability)	Internal: inadequate; test–retest: adequate only for groups with knee injuries	Face: adequate; content: cannot be assumed; construct: adequate	Responsive to change following surgery and PT; MCID: NR	Freely available; minimal burden	No patient input in development; risk of interviewer bias; multiple versions available
OKS	Pain, function; patients undergoing TKR	Patient completed	5–10 min	<5 min; manual calculation	Single score; original version 12–60 (lower scores = better outcomes); modified version 0–48 (higher scores = better outcomes)	Internal: adequate; test–retest: adequate	Face: adequate; content: adequate; construct: adequate	Responsive to change following TKR; MCID: NR	Reliable, valid, and responsive for knee OA and TKR; cross- cultural validations	Some “double-barreled” items; use of aggregate score; beware of 2 different scoring methods
WOMAC	Pain, stiffness; function; knee and hip OA	Patient- or interview-administered questionnaire (validated for in-person, telephone, and electronic use)	5–10 min	5 min; manual or computer scoring	3 subscales; range depends on version (Likert, VAS); lower scores indicate less pain,	Internal: adequate for stiffness and function, variable for pain; test–	Face: adequate; content: adequate; construct: adequate	Responsive to change following surgical and nonsurgical interventions for knee OA and	Variety of validated administration methods; validated translations into multiple languages;	Licensing and fees required; applicability of function subscale

Scale	Purpose/content	Method of administration	Respondent burden	Administrative burden	Score interpretation	Reliability evidence	Validity evidence	Ability to detect change	Strengths	Cautions
					stiffness, and functional deficits	test: variable (subscale, condition)		chondral defects; MCID for TKR and NSAID use	individual subscale scores; minimal floor and ceiling effects	items; redundant items in pain and function subscales (Rasch analysis)
Activity level										
ARS	Athletic activities; various knee disorders; participation in sport	Patient completed	<5 min	1 min; manual calculation	Single score; 0–16 (16 = more frequent participation)	Internal: NR; test–retest: adequate	Face: adequate; content: adequate; construct: adequate	Responsiveness, MCID: NR	Short and simple; adjunct to other knee function measures; generalizable across a variety of athletic and similar tasks	Recall difficulty; lack of psychometric testing
TAS	Level of sport and work participation; knee ligament injury (with Lysholm)	In-person clinician administration	3.3 min	<1 min; score corresponds to single response selected	Single score; 0–10 (higher scores = participation in higher-level activities)	Internal: N/A; test–retest: adequate (groups), less than adequate (individuals)	Face: adequate; content: cannot be assumed; construct: adequate	Responsive to change following meniscal surgery and ACL reconstruction; MCID: NR	Simple; spans work and sport/rec activities	More suited to measure within-patient change; adjustment for age and sex

*IKDC = International Knee Documentation Committee Subjective Knee Evaluation Form; MCID = minimum clinically important difference; OA = osteoarthritis; KOOS = Knee Injury and Osteoarthritis Outcome Score; ADL = activities of daily living; sport/rec = sport/recreation; QOL = quality of life; NR = not reported; KOOS-PS = Knee Injury and Osteoarthritis Outcome Score Physical Function Scale; KOS-ADL = Knee Outcome Survey Activities of Daily Living Scale; PFP = patellofemoral pain; TKR = total knee replacement; PT = physical therapy; OKS = Oxford Knee Score; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index; VAS = visual analog scale; NSAID = nonsteroidal antiinflammatory drug; ARS = Activity Rating Scale; TAS = Tegner Activity Score; N/A = not applicable; ACL = anterior cruciate ligament.