Current Concepts in Clinical Research: Anterior Cruciate Ligament Outcome Instruments

Publish date: May 4, 2018

Authors:
Molly Day, MD, ATC Kyle Hancock, MD Natalie Glass, PhD Matthew Bollier, MD

Author Affiliation | Disclosures

Authors’ Disclosure Statement: The authors report no actual or potential conflict of interest in relation to this article.

Dr. Day and Dr. Hancock are Orthopedic Surgery Residents; Dr. Glass is a Statistician; and Dr. Bollier is Congdon Professor in Orthopedic Surgery, Sports Medicine Fellowship Director, and Team Physician, Department of Orthopedics and Rehabilitation, University of Iowa Hospitals and Clinics, Iowa City, Iowa.

Address correspondence to: Matthew J. Bollier, MD, University of Iowa Hospitals and Clinics, Department of Orthopedics and Rehabilitation, 200 Hawkins Drive, Iowa City, IA 52242 (tel, 319-467-8324; fax, 319-356-8999; email, matthew-bollier@uiowa.edu).

Am J Orthop. 2018;47(5). Copyright Frontline Medical Communications Inc. 2018. All rights reserved.

Take-Home Points

- PRO instruments are widely used to capture patient perception of general health, QOL, daily function, and pain, and are an essential part of evaluation after ACL reconstruction.
- ACL outcome measures vary widely in regards to their validity, reliability, minimal clinically important difference, and applicability to specific patient populations.
- There is currently no standardized instrument universally accepted as superior following ACL reconstruction.
- In most cases, a general health outcome measure should be used in combination with a condition-specific rating scale.
- Activity rating scales, such as Marx or Tegner, should be included when evaluating patients with low-activity lifestyles.

Anterior cruciate ligament (ACL) reconstruction is one of the most common elective orthopedic procedures.\(^1\) Despite advances in surgical techniques, ACL reconstruction is associated with a lengthy recovery time, decreased performance, and increased rate of reinjury.\(^2\) Patients undergoing ACL reconstruction are often active individuals who participate in demanding activities, and accurate assessment of their recovery helps to guide recovery counseling. In addition to objective clinical outcomes measured through physical examination, patient-reported outcome (PRO) instruments add the patient’s perspective, information critical in determining a successful outcome. A variety of outcome instruments have been used and validated for patients with ACL tears. It is important for orthopedic surgeons to know the advantages and disadvantages of each outcome tool in order to
interpret clinical studies and assess postoperative patients.

Over the last 10 years, there has been an increase in the number of knee instruments and rating scales designed to measure PROs, with >54 scores designed for the ACL-deficient knee. No standardized instrument is currently universally accepted as superior following ACL reconstruction across the spectrum of patient populations. Clinicians and researchers must carefully consider an outcome instrument’s utility based on specific patient populations in which it has been evaluated. Appropriate selection of outcome measures is of fundamental importance for adequate demonstration of the efficacy and value of treatment interventions, especially in an era of healthcare reform with a focus on providing high-quality and cost-effective care.

The purpose of this review is to highlight current tools used to measure outcomes after ACL reconstruction. Current outcome measures vary widely in regards to their validity, reliability, minimal clinically important difference, and applicability to specific patient populations. We have thus identified the measures most commonly used today in studies and clinical follow-up after ACL reconstruction and their various advantages and limitations. This information may enhance the orthopedic surgeon’s understanding of what outcome measures may be utilized in clinical studies.

**Patient-Reported Outcome Instruments**

Recently, there has been a transition to increased use of PRO instruments rather than clinician-based postoperative assessment, largely due to the increasing emphasis on patient satisfaction in determining the value of an orthopedic intervention. PRO instruments are widely used to capture the patient’s perception of general health, quality of life (QOL), daily function, and pain. PRO instruments offer the benefit of allowing patients to subjectively assess their knee function during daily living and sports activities, conveying to the provider the impact of ACL reconstruction on physical, psychological, and social aspects of everyday activities. Furthermore, patient satisfaction has been shown to closely follow outcome scores related to symptoms and function. A multitude of specific knee-related PRO instruments have been developed and validated to measure outcomes after ACL reconstruction for both research and clinical purposes (Table).

**Measurement Properties**

In general, clinicians and investigators should use health-related outcome measures with established reliability, validity, patient relevance, and responsiveness for assessing the specific condition.

Reliability refers to the degree to which a measurement score is free from random error, reflecting how consistent or reproducible the instrument is when administered under the same testing conditions. Internal consistency, test-retest reliability, and measurement error are measures of reliability. Internal consistency is tested after a single administration and assesses how well items within a scale measure a single underlying dimension, represented using item-total correlation coefficients and Cronbach’s alpha. A Cronbach’s alpha of 0.70 to 0.95 is generally defined as good. Test-retest reliability is designed to appraise variation over time in stable patients and is represented using the intraclass correlation coefficient (ICC). An ICC >0.7 is considered acceptable; >0.8, good; and >0.9, excellent. An aspect of accuracy is whether the scoring system measures the full range of the disease or complaints. The incidence of minimum (floor) and maximum (ceiling) scores can be calculated for outcome scores. An instrument with low floor and ceiling effects, below 10% to 15%, is more inconclusive and can be more reliably used to measure patients at the high and low end of the scoring system.
Validity is the ability of an outcome instrument to measure what it is intended to measure. Establishing validity is complex and requires evaluation of several facets, including content validity, construct validity, and criterion validity. Content validity is a relatively subjective judgment explaining the ability of an instrument to assess the critical features of the problem. Construct validity evaluates whether the questionnaire measures what it intends to measure, and is often assessed by correlating scores from one instrument to those from other proven instruments that are already accepted as valid. Finally, criterion validity assesses the correlation between the score and a previously established “gold standard” instrument.

Responsiveness is the ability of the instrument to detect a change or identify improvement or worsening of a clinical condition over time. Most frequently, the effect size (observed change/standard deviation of baseline scores) and standardized response mean (observed change/standard deviation of change) are used as measures of responsiveness. The minimal clinically important difference of an outcome measure is the smallest change in an outcome score that corresponds to a change in patient condition.

**ACL Outcome Instruments**

**Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC LK 3.0)**

The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC LK 3.0) was developed in 1982 and is a widely used, disease-specific instrument recommended for the evaluation of treatment effects in patients with hip and knee osteoarthritis. Available in more than 80 languages, it is a self-administered, generic health status questionnaire developed to assess pain, function, and stiffness in daily living, taking respondents between 3 to 7.5 minutes for completion. Using visual analog scales, the 24 items probe the 3 subscales: pain (5 items), stiffness (2 items), and functional difficulty (17 items). Scores are calculated for each dimension, and the total score is normalized to a 100-point scale, with 0 indicating severe symptoms and 100 indicating no symptoms and higher function. The WOMAC score can also be calculated from the Knee Injury and Osteoarthritis Outcome Score (KOOS). The WOMAC questionnaire is well recognized for its good validity, reliability, and responsiveness, and is the most commonly used outcome measure for osteoarthritis. Considering its focus on older patients with osteoarthritis, it may not be appropriate for use in a young and active population.

**Knee Injury and Osteoarthritis Outcome Score (KOOS)**

The KOOS is a knee-specific questionnaire developed as an extension of the WOMAC to evaluate the functional status and QOL of patients with any type of knee injury who are at an increased risk of developing osteoarthritis. The patient-based questionnaire is available in over 30 languages and covers both the short- and long-term consequences of an injury of the knee causing traumatic damage to cartilage, ligaments, and menisci. The KOOS is 42 items graded on a 5-point Likert scale, covering 5 subscales: pain (9 items), symptoms (7 items), function in activities of daily living (17 items), function in sports/recreation (5 items), and knee-related QOL (4 items). The questionnaire is self-administered and takes about 10 minutes to complete. Scores are calculated for each dimension, and the total score is transformed to a 0 to 100 scale, with 0 representing severe knee problems and 100 representing no knee problems and better outcome. An advantage of the KOOS is that it evaluates both knee injuries and osteoarthritis; therefore, it is arguably more suitable for evaluating patients over the long-term. The KOOS has been validated for several orthopedic interventions, including ACL reconstruction and rehabilitation as well as meniscectomy and total knee replacement. Population-based reference data for the adult population
according to age and gender have also been established.\textsuperscript{20} The KOOS is increasingly utilized in clinical studies on ACL reconstruction.\textsuperscript{21-25} The questions of the WOMAC were retained so that a WOMAC score might be calculated separately and compared with the KOOS score.\textsuperscript{26}

**Patient-Reported Outcomes Information System (PROMIS)**

Since 2004, The National Institutes of Health (NIH) has funded the development of the Patient-Reported Outcome Measurement Information System (PROMIS), a set of flexible tools that reliably and validly measure PROs. The PROMIS consists of a library of question banks that has been developed and operated by a network of National Institutes of Health-funded research sites and coordinating centers and covers many different health domains including pain, fatigue, anxiety, depression, social functioning, physical functioning, and sleep. PROMIS items are developed using Item Response Theory (IRT), wherein the answer to any individual item has a known mathematical probability of predicting the test taker’s overall measurement of the specific trait being tested. This is commonly administered using computer-adaptive testing (CAT), which presents to the test taker an initial item, scores the response to that item, and from the response then presents the most informative second item, and so forth until a predefined level of precision is reached. Because the items are individually validated, they can be used alone or in any combination, a feature that distinguishes the PROMIS from traditional fixed-length PRO instruments that require the completion of an instrument in its entirety to be valid.\textsuperscript{27} In recent years, orthopedic research has been published with PROMIS physical function (PF) scores as primary outcome measures.\textsuperscript{28-30} The PF item bank includes 124 items measuring upper extremity, lower extremity, central and instrumental activities of daily living. PF can be completed as a short form (SF) with a set number of questions or utilizing CAT and evaluates self-reported function and physical activity. An advantage is its ease of use and potential to minimize test burden with very few questions, often as little as 4 items, as compared to other traditional PROMs.\textsuperscript{31}

Previously published work has demonstrated that, in patients undergoing meniscal surgery, the PROMIS PF CAT maintains construct validity and correlates well with currently used knee outcome instruments, including KOOS.\textsuperscript{28} Work by the same group looking at the performance of the PROMIS PF CAT in patients indicated for ACL reconstruction shows that the PROMIS PF CAT correlates well with other PRO instruments for patients with ACL injuries, (SF-36 PF $[r = 0.82, P < 0.01]$, KOOS Sport $[r = 0.70, P < 0.01]$, KOOS ADL $[r = 0.74, P < 0.01]$), does not have floor or ceiling effects in this relatively young and healthy population, and has a low test burden.\textsuperscript{32,33} Papuga and colleagues\textsuperscript{33} also compared the International Knee Documentation Committee (IKDC) and PROMIS PF CAT on 106 subjects after ACL reconstruction and found good correlation.

**Quality of Life Outcome Measure for ACL Deficiency (ACL-QOL)**

The ACL-QOL Score was developed in 1998 as a disease-specific measure for patients with chronic ACL deficiency.\textsuperscript{34} This scale consists of 32 separate items in 31 visual analog questions regarding symptoms and physical complaints, work-related concerns, recreational activities and sport participation or competition, lifestyle, and social and emotional health status relating to the knee. The raw score is transformed into a 0- to 100-point scale, with higher scores indicating a better outcome. The scale is valid, reliable, and responsive for patients with ACL insufficiency,\textsuperscript{35,36} and is not applicable to other disorders of the knee. We recommend the ACL-QOL questionnaire be used in conjunction with other currently available objective and functional outcome measures.

**Cincinnati Knee Rating System**

The Cincinnati Knee Rating System (CKRS) was first described in 1983 and was modified to include occupational
activities, athletic activities, symptoms, and functional limitations. There are 11 components, measuring symptoms and disability in sports activity, activities of daily living function, occupational rating, as well as sections that measure physical examination, laxity of the knee, and radiographic evidence of degenerative joint disease. The measure is scored on a 100-point scale, with higher scores indicating better outcomes. Scores have been shown to be lower as compared with other outcome measures assessing the same clinical condition. Barber-Westin and colleagues confirmed the reliability, validity, and responsiveness of the CKRS by testing 350 subjects with and without knee ligament injuries. In 2001, Marx tested the CKRS subjective form for reliability, validity, and responsiveness and found it to be acceptable for clinical research.

**Lysholm Knee Score**

The Lysholm Knee Score was published in 1982 and modified in 1985, consisting of an 8-question survey that evaluates outcomes after knee ligament surgery. Items include pain, instability, locking, squatting, limping, support usage, swelling, and stair-climbing ability, with pain and instability carrying the highest weight. It is scored on a scale of 0 to 100, with high scores indicating higher functioning and fewer symptoms. It has been validated in patients with ACL injuries and meniscal injuries. Although it is widely used to measure outcomes after ACL reconstruction, it has received criticism in the evaluation of patients with other knee conditions. The main advantage of the Lysholm Knee Score is its ability to note changes in activity in the same patient across different time periods (responsiveness). A limitation of the Lysholm Knee Score is that it does not measure the domains of functioning in daily activities, sports, and recreational activities. The Lysholm scoring system’s test-retest reliability and construct validity have been evaluated, although there has been some concern regarding a ceiling effect and its validity, sensitivity, and reliability has been questioned. Therefore, it is advised that this score be used in conjunction with other PRO scores.

**International Knee Documentation Committee (IKDC) Subjective Knee Form**

In 1987, members of the European Society for Knee Surgery and Arthroscopy and the American Orthopaedic Society for Sports Medicine formed the IKDC to develop a standardized method for evaluating knee injuries and treatment. The IKDC Subjective Knee Evaluation Form was initially published in 1993, and in 2001 the form was revised by the American Orthopaedic Society for Sports Medicine to become a knee-specific assessment tool rather than a disease or condition-specific tool. The IKDC subjective form is an 18-question, knee-specific survey designed to detect improvement or deterioration in symptoms, function, and ability to participate in sports activities experienced by patients following knee surgery or other interventions. The individual items are summed and transformed into a 0- to 100-point scale, with high scores representing higher levels of function and minimal symptoms. The IKDC is utilized to assess a variety of knee conditions including ligament, meniscus, articular cartilage, osteoarthritis, and patellofemoral pain. Thus, this form can be used to assess any condition involving the knee and allow comparison between groups with different diagnoses. The IKDC has been validated for an ACL reconstruction population, has been used to assess outcomes in recent clinical studies on ACL reconstruction, and is one of the most frequently used measures for patients with ACL deficiency. The validity, responsiveness, and reliability of the IKDC subjective form has been confirmed for both adult and adolescent populations.

**Tegner Activity Scores**

The Tegner activity score was developed in 1985 and was designed to provide an objective value for a patient’s
activity level. This scale was developed to complement the Lysholm score. It consists of 1 sport-specific activity level question on a 0 to 10 scale that evaluates an individual’s ability to compete in a sporting activity. Scores between 1 and 5 represent work or recreational sports. Scores >5 represent higher-level recreational and competitive sports. The Tegner activity score is one of the most widely used activity scoring systems for patients with knee disorders, commonly utilized with the Lysholm Knee Score. One disadvantage of the Tegner activity score is that it relates to specific sports rather than functional activities, which limits its generalizability. We are not aware of any studies documenting the reliability or validity of this instrument.

Marx Activity Rating Scale

The Marx activity rating scale was developed to be utilized with other knee rating scales and outcome measures as an activity assessment. In contrast to the Tegner activity score, the Marx activity rating scale measures function rather than sport-specific activity. The scale is a short, patient-based activity assessment that consists of a 4-question survey evaluating patients’ knee health by recording the frequency and intensity of participation in a sporting activity. Questions are scored from 0 to 4 on the basis of how often the activity is performed. The 4 sections of the Marx scale that are rated include running, cutting, decelerating, and pivoting. This scale has been validated in patients with ACL injuries, chondromalacia patellae, and meniscal lesions. Acceptable ceiling effects of 3% and floor effects of 8% were noted in the study of ACL-injured patients.

American Academy of Orthopaedic Surgeons (AAOS) Sports Knee Scale

The American Academy of Orthopaedic Surgeons (AAOS) Sports Knee Rating Scale consists of 5 parts and 23 items, including a section addressing stiffness, swelling, pain and function (7 questions), locking/catching (4 questions), giving way (4 questions), limitations of activity (4 questions), and pain with activity (4 questions). Items may be dropped if patients select particular responses, which can lead to difficulties when using the survey. This scoring system has been found to be satisfactory when all subscales were combined and the mean was calculated.

Discussion

PRO measures play an increasingly important role in the measurement of success and impact of health care services. Specifically, for ACL reconstruction, patient satisfaction is key for demonstrating the value of operative or other interventions. Selecting a suitable outcome measurement tool can be daunting, as it can be difficult to ascertain which outcome measures are appropriate for the patient or disorder in question. As there is currently no instrument that is universally superior in the evaluation of ACL outcomes, clinicians must consider the specific patient population in which the outcome instrument has been evaluated. Investigators should also use instruments with reported minimal clinically important differences so that variation in scores can be interpreted as either clinically significant or not. When choosing which outcome instrument to use, there is rarely a single appropriate rating system that is entirely comprehensive. In most cases, a general health outcome measure should be used in combination with a condition-specific rating scale. Activity rating scales, such as Marx or Tegner, should be included, especially when evaluating patients with low-activity lifestyles.
Conclusion

There are a number of reliable, valid, and responsive outcome measures that can be utilized to evaluate outcomes following ACL reconstruction in an array of patient populations. Outcome measures should be relevant to patients, easy to use, reliable, valid, and responsive to change. By increasing familiarity with these outcome measures, orthopedic surgeons and investigators can develop better studies, interpret data, and implement findings in practice with sound and informed judgment. Future research should focus on identifying the most relevant outcome metrics for assessing function following ACL reconstruction.

This paper will be judged for the Resident Writer’s Award.

Key Info

Figures/Tables

Table. ACL Outcome Measures

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>Condition/Intervention</th>
<th>Measures</th>
<th>Internal Consistency (Cronbach’s a)</th>
<th>Test-Retest Reliability</th>
<th>Minimal Clinically Important Difference</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAOS Sports Knee Scale</td>
<td>Many Knee</td>
<td>Stiffness, swelling, pain/function, locking/catching, living way, limitation of activity, pain with activity</td>
<td>0.86-0.95</td>
<td>0.68-0.96</td>
<td>Unknown</td>
<td>59, 60</td>
</tr>
<tr>
<td>ACL-QOL</td>
<td>Chronic ACL deficiency</td>
<td>Physical complaints, work, recreation and sports competition, lifestyle, social and emotional functioning</td>
<td>0.93-0.98</td>
<td>0% average error</td>
<td>Unknown</td>
<td>35, 36</td>
</tr>
<tr>
<td>Cincinnati Knee Rating System</td>
<td>ACL</td>
<td>Symptoms, daily and sports activities, physical examination, stability, radiographs, functional testing</td>
<td>0.80-0.97</td>
<td></td>
<td>14 points (6 months), 26 points (12 months)</td>
<td>39, 40, 47, 52</td>
</tr>
<tr>
<td>IKDC (Subjective Knee Form)</td>
<td>ACL</td>
<td>Symptoms, function, sports activity</td>
<td>0.92</td>
<td>0.91-0.93</td>
<td>13.5 points; 6.3 at 6 months, 16.7 at 12 months</td>
<td>48, 52, 54</td>
</tr>
<tr>
<td>KOOS</td>
<td>ACL</td>
<td>Pain, symptoms, activities of daily living, sporting/recreation, knee-related quality of life</td>
<td>0.71-0.95</td>
<td>0.75-0.93</td>
<td>0-10 points</td>
<td>57</td>
</tr>
<tr>
<td>Lysholm</td>
<td>ACL</td>
<td>Pain, instability, locking, squatting, limp, support, swelling, stair-climbing</td>
<td>0.72</td>
<td>0.94</td>
<td>0.9</td>
<td>46, 47, 53</td>
</tr>
<tr>
<td>Marx</td>
<td>Healthy patients</td>
<td>Activity level</td>
<td>0.87</td>
<td>0.97</td>
<td>Unknown</td>
<td>52, 56, 57</td>
</tr>
<tr>
<td>Tegner</td>
<td>ACL</td>
<td>Activity level</td>
<td>0.81</td>
<td>0.82</td>
<td>1</td>
<td>55, 56</td>
</tr>
<tr>
<td>PROMIS (PF CAT)</td>
<td>Many lower extremity orthopedic conditions</td>
<td>Lower extremity function, central body function, activities of daily living</td>
<td>0.98</td>
<td>0.96-0.99</td>
<td>12% baseline score or 9% max score; 9-12 points</td>
<td>50, 31</td>
</tr>
<tr>
<td>WOMAC</td>
<td>Hip/knee OA</td>
<td>Physical function, pain, stiffness</td>
<td>0.81-0.95</td>
<td>0.80-0.92</td>
<td>12% baseline score or 9% max score; 9-12 points</td>
<td>53, 14</td>
</tr>
</tbody>
</table>

Abbreviations: AAOS, American Academy of Orthopaedic Surgeons; ACL, anterior cruciate ligament; ACL-QOL, anterior cruciate ligament quality of life score; CAT, computer-adapting testing; IKDC, International Knee Documentation Committee; KOOS, Knee Injury and Osteoarthritis Outcome Score; OA, osteoarthritis; PF, physical function; PROMIS, Patient-Reported Outcome Measurement Information System; Ref, references; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.
References


13. Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of


25. Duffee A, Magnussen RA, Pedroza AD, Flanigan DC; MOON Group, Kaeding CC. Transtibial ACL


**Multimedia**

**Product Guide**

- STRATAFIX™ Symmetric PDS™ Plus Knotless Tissue Control Device
- STRATAFIX™ Spiral Knotless Tissue Control Device
• BioComposite SwiveLock Anchor
• BioComposite SwiveLock C, with White/Black TigerTape™ Loop

Citation

Publish date: May 4, 2018

Molly Day, MD, ATC