Evaluation of Functional Knee Instability in Sports: A Bibliometric Review of Patient-Reported Outcome Measures and Their Measurement Properties

Master's Thesis

Ian Filipe Champier de Oliveira

This thesis was completed under the supervision of:

Dr. Pedro Duarte-Mendes, Department of Sports and Well-being, Polytechnic Institute of Castelo Branco

Dr. Diogo Monteiro, School of Education and Social Sciences, Polytechnic of Leiria

Leiria, May 2026

Master's degree in Exercise Prescription and Health Promotion

SCHOOL OF EDUCATION AND SOCIAL SCIENCES

POLYTECHNIC OF LEIRIA

ACKNOWLEDGEMENTS

Great works are never built alone, but through the strength of those who stand beside us. For that reason, I am deeply grateful for all the support, encouragement, and help I received throughout the completion of this master's thesis.

First, I would like to express my heartfelt gratitude to my supervisors, Professor Pedro Duarte-Mendes and Professor Diogo Monteiro. I sincerely thank them for their guidance, knowledge, and constant willingness to help whenever needed. To Professor Pedro Duarte-Mendes, I am especially grateful for his friendship, availability, and the many long conversations that accompanied this journey.

I would also like to thank my mother, for always encouraging me to pursue my dreams, and my brother, for being an example of perseverance and showing me that it is never too late to study. To my family, friends, TUSALD, and EACB, thank you for your unconditional support.

A special dedication goes to my father, grandparents, godfather, and my dog, whom I will forever hold in my heart and memory.

Finally, to my fiancée, thank you for your patience, motivation, and love throughout this journey.

RESUMO

Enquadramento: A instabilidade do joelho é uma queixa clinicamente relevante no desporto; contudo, a avaliação dos outcomes recorre frequentemente a medidas de resultado reportadas pelo doente (*patient-reported outcome measures* — PROMs), cujos constructos podem não estar especificamente adaptados à instabilidade funcional em tarefas desportivas.

Objetivo: Identificar os PROMs mais influentes utilizados para avaliar a instabilidade funcional do joelho em contextos desportivos e descrever o seu perfil bibliométrico, juntamente com uma síntese das propriedades de medida reportadas, informada pela abordagem COSMIN.

Métodos: Foi realizada uma revisão bibliométrica nas bases de dados Web of Science, Scopus, PubMed e SportDiscus, desde a sua criação até agosto de 2025. Foram elegíveis estudos originais com atletas ou indivíduos fisicamente ativos em contexto desportivo que reportassem pelo menos um PROM específico do joelho como outcome. Os registos foram triados segundo critérios definidos *a priori*. A extração incluiu características da publicação, indicadores de citação, métricas das revistas, país do primeiro autor e características dos instrumentos. As propriedades de medida e o risco de viés foram avaliados através da abordagem COSMIN Risk of Bias.

Resultados: Dos 10 059 registos identificados, 20 documentos cumpriram os critérios de inclusão e foram publicados em 13 revistas, predominantemente nas áreas da ortopedia e medicina desportiva. A produção científica aumentou após 2017, com 17 dos 20 documentos publicados entre 2017 e 2025. A maioria das revistas era Q1, e o fator de impacto médio ponderado de 2024 foi 4,22. Onze países estiveram representados de acordo com a afiliação do primeiro autor; os Estados Unidos corresponderam à maior proporção de publicações (6/20) e ao maior número total de citações. A distribuição das citações concentrou-se num pequeno conjunto de PROMs: o IKDC-SKF foi o instrumento mais citado, com 160 citações, seguido pelo Lysholm e pelo KOS-SAS. Medidas que avaliam a exposição à atividade (Tegner) e a prontidão psicológica (ACL-RSI) também estiveram representadas, embora com menor proeminência. A fiabilidade foi a propriedade de medida mais consistentemente suportada, enquanto a evidência relativa à validade de conteúdo, erro de medição, validade transcultural e responsividade permaneceu limitada ou reportada de forma inconsistente.

Conclusão: A investigação sobre resultados relacionados com a instabilidade do joelho no desporto é bibliometricamente dominada por PROMs clássicos de função do joelho, particularmente o IKDC-SKF e o Lysholm, num corpus publicado sobretudo em revistas de elevado impacto e com um aumento marcado da produção científica após 2017. Contudo, a proeminência bibliométrica não implica necessariamente especificidade de constructo nem adequação psicométrica abrangente para a instabilidade funcional relacionada com o desporto. Estes resultados reforçam a necessidade de uma seleção criteriosa dos instrumentos e, quando apropriado, da utilização complementar de PROMs dirigidos à exposição à atividade, prontidão psicológica, impacto na qualidade de vida e exigências do retorno ao desporto.

Palavras chave

Instabilidade funcional do joelho, medidas de resultado reportadas pelo doente, revisão bibliométrica, COSMIN, retorno ao desporto.

ABSTRACT

Background: Knee instability is a clinically relevant complaint in sport; however, outcome assessment frequently relies on patient-reported outcome measures (PROMs), whose constructs may not be specifically tailored to functional instability during sport-specific tasks.

Objective: To identify the most influential PROMs used to assess functional knee instability in sporting contexts and to describe their bibliometric profile, together with a COSMIN-informed synthesis of the reported measurement properties.

Methods: A bibliometric review was conducted in Web of Science, Scopus, PubMed, and SportDiscus, from database inception to August 2025. Original studies involving athletes or physically active individuals in a sporting context and reporting at least one knee-specific PROM as an outcome were eligible. Records were screened according to *a priori* criteria. Data extraction included publication characteristics, citation indicators, journal metrics, first-author country, and instrument-level characteristics. Measurement properties and risk of bias were assessed using the COSMIN Risk of Bias approach.

Results: From 10,059 identified records, 20 documents met the inclusion criteria and were published across 13 journals, mainly in orthopaedics and sports medicine. Scientific output increased after 2017, with 17 of the 20 documents published between 2017 and 2025. Most journals were Q1, and the weighted mean 2024 impact factor was 4.22. Eleven countries were represented according to first-author affiliation; the United States accounted for the largest proportion of publications (6/20) and the highest total number of citations. Citation distribution was concentrated in a small set of PROMs: the IKDC-SKF was the most cited instrument, with 160 citations, followed by Lysholm and KOS-SAS. Measures assessing activity exposure (Tegner) and psychological readiness (ACL-RSI) were represented, although less prominently. Reliability was the most consistently supported measurement property, whereas evidence for content validity, measurement error, cross-cultural validity, and responsiveness remained limited or inconsistently reported.

Conclusion: Research on knee instability outcomes in sport is bibliometrically dominated by legacy knee function PROMs, particularly the IKDC-SKF and Lysholm, within a corpus published mainly in high-impact journals and showing a marked increase in scientific output after 2017. However, bibliometric prominence does not necessarily imply construct specificity or comprehensive psychometric adequacy for sport-related functional instability. These findings reinforce the need for careful instrument selection and, where appropriate, complementary PROMs targeting activity exposure, psychological readiness, quality-of-life impact, and return-to-sport demands.

Keywords

Knee instability; patient-reported outcome measures; bibliometric analysis; sports medicine; psychometric properties.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

Abbreviation	Meaning	Abbreviation	Meaning
ACL	Anterior Cruciate Ligament	**ACLR**	Anterior Cruciate Ligament Reconstruction
ACL-QOL	Anterior Cruciate Ligament Quality of Life	**ACL-RSI**	Anterior Cruciate Ligament Return to Sport after Injury
AKPS	Anterior Knee Pain Scale	**AP**	Antero-Posterior
AUC	Area Under the Curve	**CKRS**	Cincinnati Knee Rating System
COSMIN	COnsensus-based Standards for the selection of health Measurement INstruments	**IF**	Impact Factor
ICC	Intraclass Correlation Coefficient	**IKDC-SKF**	International Knee Documentation Committee Subjective Knee Form
KOOS	Knee injury and Osteoarthritis Outcome Score	**KOS-ADLS**	Knee Outcome Survey Activities of Daily Living Scale
KOS-SAS	Knee Outcome Survey Sports Activities Scale	**LCL**	Lateral Collateral Ligament
MCL	Medial Collateral Ligament	**MCID**	Minimal Clinically Important Difference
MDC	Minimal Detectable Change	**MIC**	Minimal Important Change
MPFL	Medial Patellofemoral Ligament	**PA**	Postero-Anterior
PCL	Posterior Cruciate Ligament	**Pedi-IKDC**	Paediatric International Knee Documentation Committee Subjective Knee Form

Abbreviation	Meaning	Abbreviation	Meaning
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses	**PROM**	Patient-Reported Outcome Measure
PROMs	Patient-Reported Outcome Measures	**Q1**	First Quartile
Q2	Second Quartile	**RMSEA**	Root Mean Square Error of Approximation
RTS	Return to Sport	**SDC**	Smallest Detectable Change
SEM	Standard Error of Measurement	**SJR**	SCImago Journal Rank
SRM	Standardised Response Mean	**TLS**	Total Link Strength
VOSviewer	Visualisation of Similarities viewer		

1. INTRODUCTION

The knee joint is a complex anatomical and biomechanical system in which osseous geometry, menisci, capsuloligamentous structures and periarticular muscles interact to provide both mobility and dynamic stability. Because tibiofemoral motion occurs across multiple planes, knee stability cannot be reduced to a single passive restraint or to an isolated mechanical parameter. Rather, it reflects the integrated contribution of primary and secondary stabilisers that constrain anterior–posterior, varus–valgus and rotational motion while allowing the joint to tolerate sport-specific loading demands (summarized in Supplementary Table 1)(Abulhasan & Grey, 2017; Sundemo et al., 2016; Zlotnicki et al., 2016).

From the patient's perspective, knee instability is commonly reported as a sensation of shifting, buckling or giving way (Schmitt & Rudolph, 2008). Clinically, true mechanical instability has been described as excessive joint translation beyond the limits imposed by ligamentous, muscular and articular restraints, potentially resulting in partial or complete loss of joint congruence (Cross, 1996). Rotatory knee instability, in turn, refers to abnormal three-dimensional motion involving complex interactions between ligamentous structures, menisci, osseous alignment and neuromuscular control (Hughes et al., 2019). These definitions are important, but they do not fully capture the sport-context construct addressed in the present review. Because the term functional knee instability is used heterogeneously across clinical and sports-medicine literature, this review adopts an operational definition centred on the athlete's self-reported experience of recurrent knee shifting, buckling, or giving way during functional or sport-related activities. This definition distinguishes perceived functional instability from purely mechanical laxity or structural subluxation, while recognising that these phenomena may coexist. Accordingly, functional knee instability is conceptualised here as a task-dependent construct that may emerge most clearly during high-velocity, multiplanar and reactive sporting actions, where passive restraints, neuromuscular control, confidence and sport-specific demands interact.

The scientific relevance of this construct is supported by the epidemiological burden of knee injuries in sport and by their implications for function, performance and return-to-sport decision-making. In a prospective cohort study of professional football players conducted over seven consecutive seasons, injuries associated with knee instability ranked third among all injuries sustained (Ekstrand et al., 2011). In another seven-year longitudinal study, knee injury accounted for 15–50% of all injuries across 12 different recreational sports practiced by youth men and women (de Loës et al., 2000). Such heterogeneity suggests that knee instability should not be conceptualised as a uniform clinical event; pivoting, cutting, jumping and contact or collision demands may expose different mechanical and neuromuscular constraints than linear or lower-impact activities. Accordingly, the assessment of functional knee instability in athletes requires instruments that are sensitive not only to symptoms, but also to the ecological demands of sport participation and performance.

This measurement problem remains unresolved. Objective approaches such as passive laxity testing, instrumented arthrometry, postural balance tasks, gait analysis, perturbation paradigms, kinematic assessment and electromyographic outcomes can provide valuable mechanistic information. However, these approaches do not yet constitute a universally accepted standard for quantifying knee instability, and many are obtained under laboratory or non-sport-specific conditions that may insufficiently reproduce the perceptual and motor demands of sport. Consequently, the central limitation is not simply whether an instrument is objective or subjective, but whether it has ecological and construct validity for capturing instability as experienced during sport. This is particularly relevant because perceived instability is multifactorial and may reflect interactions between passive restraint deficiency, strength deficits, altered muscle activation, proprioceptive impairment, pain, fear, psychological readiness and reduced functional confidence (Chaharmahali et al., 2021; Schrijvers et al., 2019; Sharma et al., 2015; van der Esch et al., 2016).

Despite the extensive PROM landscape for knee conditions, the proliferation of instruments does not necessarily equate to construct-specific measurement of functional knee instability in sporting contexts. Under this operational definition, PROMs are indispensable because they can capture the athlete’s perceived capacity to trust and control the knee during activity. However, a recent scoping meta-review of systematic reviews in young adults identified 80 knee-related PROMs, highlighting both the breadth of available instruments and the heterogeneity in what is being measured and how measurement properties are reported (Bausch et al., 2025). Moreover, commonly used knee-specific PROMs often contain substantial overlap in item content, including items related to symptoms, sport/recreation function, quality of life, psychological readiness and “giving way”, suggesting that apparent variety may partly reflect the re-packaging of similar domains rather than targeted capture of sport-context functional instability (Collins et al., 2011; Jansson et al., 2021). Crucially, in widely adopted instruments such as the IKDC, “giving way” is typically embedded within broader symptom/function domains rather than operationalised as a comprehensive, sport-demand-sensitive construct (Collins et al., 2011). This is a non-trivial limitation because perceived instability is increasingly understood as multifactorial, encompassing not only tissue-level laxity, but also strength, altered muscle activation patterns, proprioception, pain, confidence and neuromuscular control, dimensions that may manifest most clearly during high-velocity, multiplanar and reactive sporting tasks (Chaharmahali et al., 2021). In parallel, return-to-sport decision-making after ligament injury is explicitly described as a multidimensional process that should incorporate functional assessment and psychological readiness, reinforcing that instability in athletes cannot be assumed to be fully represented by generic symptom/function scores alone (Diermeier et al., 2021; Webster et al., 2008).

This distinction is consistent with contemporary COSMIN guidance, which places content validity at the centre of PROM evaluation. Relevance, comprehensiveness and comprehensibility for the target construct and population are prerequisites for meaningful interpretation of measurement scores (Mokkink et al., 2024; Terwee et al., 2018). For this reason, citation frequency should not be interpreted as evidence of measurement quality. Rather, citation analysis is useful for identifying which PROMs have shaped the field, how they have diffused across journals, countries, populations and sport-related research areas, and which instruments are most likely to influence clinical and research decision-making. Bibliometric influence must therefore be interpreted alongside, not instead of, evidence on measurement properties. This integrative approach is necessary because a highly cited instrument may reflect popularity, historical precedence or broad applicability, while COSMIN-informed appraisal can determine whether its content and psychometric evidence are adequate for assessing functional knee instability in sporting contexts.

Accordingly, the present review adopts an integrative bibliometric and measurement-property perspective. It aims to identify the most influential PROMs used to assess functional knee instability in sporting contexts and to describe their bibliometric profile alongside a COSMIN-informed synthesis of reported measurement properties. By combining citation-based mapping with construct-focused interpretation of PROM content and psychometric evidence, this review seeks to clarify which instruments dominate the literature, what they actually measure, and whether their use is defensible for capturing functional knee instability in athletes and physically active populations.

2. MATERIALS AND METHODS

2.1 ELIGIBILITY CRITERIA

Eligible records were original research articles published in peer-reviewed journals and indexed in the selected databases, reporting data from human participants and providing at least title, abstract and basic bibliographic information. Studies were considered for inclusion when they primarily addressed structural, traumatic or perceived knee joint instability, such as anterior or posterior cruciate ligament injury, collateral or posterolateral corner instability, multiligament knee injuries, or patellofemoral/patellar instability, with instability (for example, “giving way”, buckling, subluxation or joint laxity) treated as a core clinical construct rather than a secondary or incidental feature.

The population of interest comprised athletes and physically active individuals involved in organised or regular sports participation, including professional, semiprofessional, federated or recreational athletes, typically corresponding to a Tegner Activity Scale score of 5 or higher when reported, or an equivalent description of sports engagement. To be eligible, studies had to include at least one knee-specific patient-reported outcome measure (PROM), in its original, adapted or shortened version, as an outcome.

Knee joint instability was operationally defined in line with established clinical terminology as a subjective complaint of the knee “giving way”, buckling or subluxing due to pathological laxity or abnormal joint motion, rather than pain or general functional limitation alone (Cross, 1996; Hughes et al., 2019; Schmitt & Rudolph, 2008).

Studies were excluded when they focused exclusively on knee pain, osteoarthritis, meniscal pathology, general function or performance outcomes without a clear reference to knee joint instability as defined above, or when “instability” referred to concepts unrelated to joint laxity or perceived knee giving way, such as postural or gait instability. Articles were also excluded if the study population did not clearly reflect a sports context, if none of the knee-specific PROMs were used, or if the publication type was a review, editorial, commentary, conference abstract without a corresponding full-text article, animal/cadaver study, or methodological paper not reporting original development, validation, cross-cultural adaptation or measurement-property data for a knee-specific PROM. Duplicate records across databases were removed, retaining the most complete entry. When, after title and abstract screening and, where necessary, full-text assessment, essential information regarding sports context, type of instability or PROMs used could not be ascertained, the study was excluded due to insufficient information or being outside the scope of the review.

2.2 Search Strategy

Following the bibliometric approach described by (Clancy et al., 2017), we analysed the citation profile of the PROMs included in this review. For each questionnaire, we extracted the total number of citations, annual citation rate, citing journals, number of citing manuscripts and journal-level indicators. This procedure allowed us to identify the most cited instruments, the main journals and articles associated with their use, and the relative influence of each PROM within the literature on functional knee instability in sporting contexts. Four electronic databases were searched from database inception to August 2025, namely Web of Science, Scopus, PubMed and SportDiscus. The search strategy included both MeSH terms, when available, and free-text descriptors, adapted according to the syntax and indexing structure of each database.

The search strategy combined terms related to the knee joint, instability, sport participation, measurement properties and patient-reported instruments. Boolean operators "AND" and "OR" were used to combine the following search concepts: ("knee" OR "knee joint" OR "knee injury" OR "functional instability") AND (sport* OR exercise* OR training) AND ("psychometric properties" OR "reliability" OR "validity" OR "validation studies" OR "internal consistency") AND (survey OR questionnaire* OR instrument* OR measur* OR scale OR "subjective evaluation" OR assess* OR tool OR "self-report*").

The full search strategy was adapted for each database and is reported in the supplementary material. An example of the PubMed search strategy is provided in Supplementary Table 2.

2.3 Study Selection Process

Following completion of the database searches, all retrieved records were exported to EndNote v21.1 for reference management and initial duplicate removal. The remaining records were then checked manually in Microsoft Excel to identify residual duplicate entries before title and abstract screening. The final number of duplicate records removed before screening was reported in the PRISMA flow diagram. The study selection process was conducted sequentially and documented to support construction of the PRISMA flow diagram. In the first stage, titles and abstracts were screened to identify records potentially related to knee joint instability, sport or regular physical activity, and the use of knee-specific patient-reported outcome measures. Records clearly unrelated to the review scope were excluded at this stage. In the second stage, records considered potentially eligible, or insufficiently clear at title and abstract level, were retrieved for full-text assessment. Full-text articles were then assessed against the predefined eligibility criteria, including study population, sport-related context, knee instability construct, publication type, and PROM use. When a study met more than one exclusion criterion, the primary reason for exclusion was recorded according to the criterion most directly responsible for ineligibility. Exclusion reasons at the full-text stage were grouped into the following categories: absence of a clear sport-related population; absence of knee joint instability as a core construct; absence of a knee-specific PROM; publication type outside the scope of the review; non-human, cadaveric or ex vivo study design; methodological paper without original PROM development, validation, cross-cultural adaptation or measurement-property data; and insufficient information to determine eligibility.

Two reviewers independently conducted the screening and selection process. Disagreements were resolved through discussion and consensus, and when necessary, a third reviewer was consulted. The number of records identified, duplicates removed, records screened, reports assessed for eligibility, reports excluded with reasons, and studies included in the final synthesis were recorded and reported in the PRISMA flow diagram.

2.4 Data Extraction

Data extraction was performed independently by two reviewers using a predefined extraction form developed specifically for this review. The form included bibliographic variables, study-level characteristics, PROM-related information, citation indicators, journal-level metrics and reported measurement-property data. Disagreements were resolved by discussion and consensus, and when necessary, a third reviewer was consulted. Before full data extraction, the form was pilot-tested in a subset of eligible studies including different PROMs and study designs. This pilot phase was used to refine variable definitions, clarify ambiguous fields and improve consistency between reviewers.

From each eligible primary study, the following variables were extracted: scale name; first author; publication date; country of the first author; journal of publication; total number of citations of the journal; detailed scale information, including test–retest reliability, number of items, subscales, response options per item and scoring/rating procedures; number of citations of the article; citations per year; impact factor of the original journal; journal quartile; number of journals that cited the instrument; number of manuscripts that cited the instrument/measures; and the weighted average impact factor of citing journals. Citation counts were retrieved in August 2025 to ensure consistency across studies. Journal-level indicators, including 2024 Journal Impact Factor and SJR quartile, were obtained from SCImago in December 2025. EndNote v21.1 was used for reference management, duplicate removal and to assist data extraction.

2.5 Bibliometric Analysis and Network Visualisation

Bibliometric data were prepared for descriptive and network analyses. Citation indicators were analysed through the number of citations, annual citations, journal impact indicators, journal quartile, number of citing journals, number of citing manuscripts and weighted average impact factor of citing journals. PROMs were ranked according to citation frequency and interpreted together with the reported measurement properties.

Bibliometric networks, including co-authorship, country, organisation and author keyword co-occurrence maps, were constructed and visualised using VOSviewer. Full-counting or fractional-counting methods were applied according to the type of analysis, and minimum occurrence thresholds were defined to improve the interpretation of the resulting maps. Network visualisations were used to identify collaboration patterns, geographical distribution, organisational contribution and thematic clusters within the literature on PROMs used to assess functional knee instability in sporting contexts.

2.6 COSMIN Risk of Bias approach

The methodological quality of studies reporting measurement properties of the included PROMs was appraised using the COSMIN Risk of Bias checklist for systematic reviews of patient-reported outcome measures. The assessment focused on the measurement-property domains reported for each instrument, including structural validity, internal consistency, reliability, measurement error, hypothesis testing for construct validity, responsiveness and cross-cultural validity/measurement invariance. Content validity evidence, when available, was considered with particular attention to the relevance, comprehensiveness and comprehensibility of each PROM for assessing functional knee instability in sporting contexts.

The COSMIN Risk of Bias assessment was performed by one reviewer and checked by a second reviewer. Disagreements or uncertainties were resolved through discussion and, where necessary, consultation with a third reviewer. Each measurement property was rated according to the COSMIN four-point rating system, namely “very good”, “adequate”, “doubtful” or “inadequate”, with the final methodological quality judgement determined by the lowest rating within each domain.

Because this review adopts an integrative bibliometric and measurement-property perspective, the COSMIN Risk of Bias assessment was used to contextualise the methodological credibility of the reported psychometric evidence rather than to replace the citation-based bibliometric analysis. Accordingly, citation frequency was interpreted as an indicator of influence and diffusion, whereas COSMIN ratings were interpreted as indicators of the methodological quality of the available measurement-property evidence.

3. Results

Starting from 10,059 records retrieved, title and abstract screening identified 81 studies of potential interest. Full-text evaluation, guided by a priori eligibility criteria, led to the identification of the most frequently referenced measures for assessing knee instability, which were included in the critical analysis of this review (Figure 1).

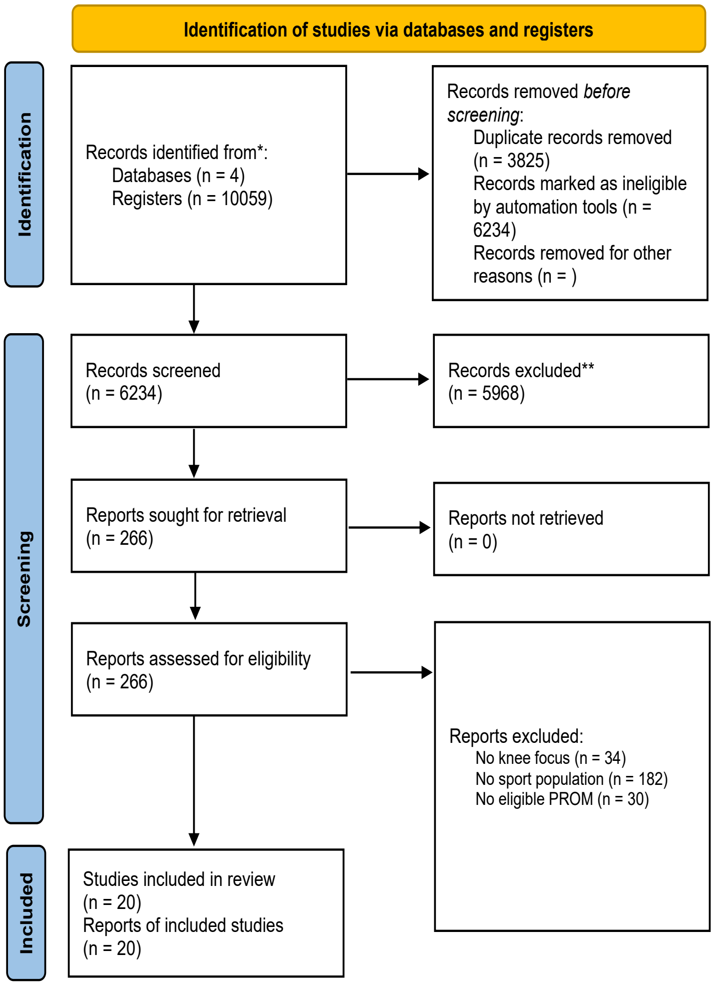


Figure 1. Diagram illustrating each phase of the search and selection process (Page et al., 2021).

3.1. Annual publication trends

A total of 20 documents met the inclusion criteria and were included in the bibliometric analysis. Figure 2 displays the annual distribution of publications, with the x-axis representing the publication year and the y-axis the number of documents. The earliest articles were published in 2001, 2002 and 2012, each year contributing a single document. No publications were identified between 2003 and 2011, nor in 2013, 2014, 2015, 2016, 2018 or 2020. From 2017 onwards, however, a clear increase in research activity is evident, with at least one document published in every recorded year. Peaks of four publications were observed in 2017 and 2019, followed by two documents in both 2021 and 2022, and single documents in 2023 and 2024. The year 2025 shows a secondary rise with three documents. Overall, this pattern reflects a marked intensification of scientific output over time, with a more than fivefold increase in the number of publications when comparing the early period (2001–2012, three documents) with the more recent years (2017–2025, 17 documents).

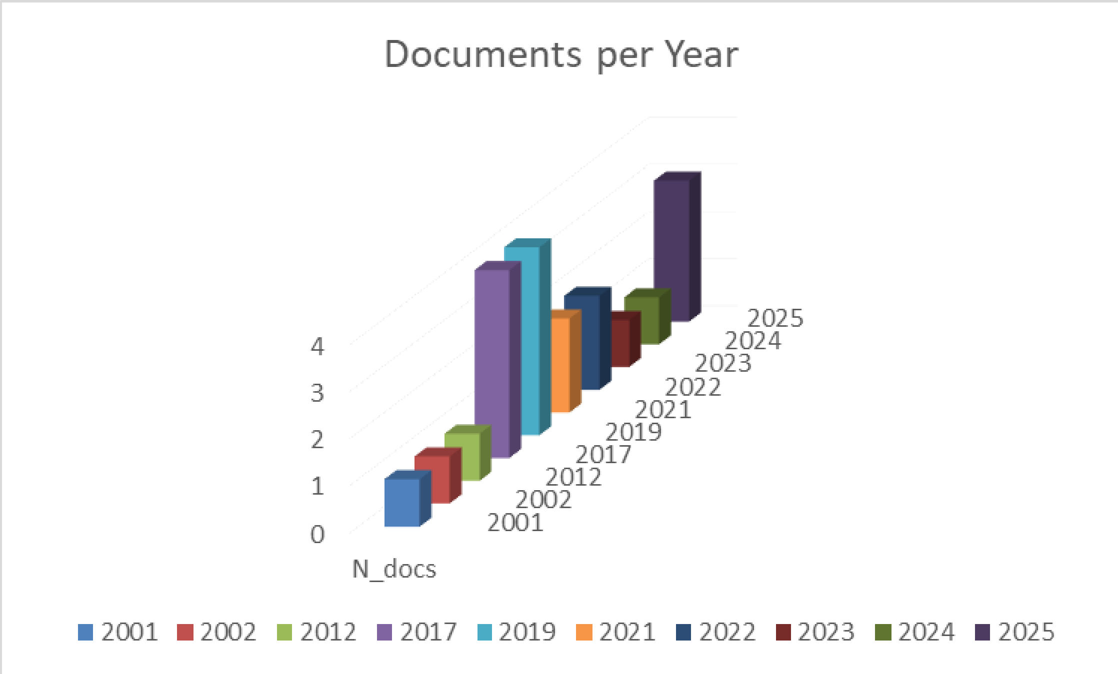


Figure 2. Documents published by Year

3.2. Subject categories (SCImago)

Because SCImago subject categories are not mutually exclusive, the same journal and article could contribute to more than one category; therefore, the percentages reported in this section should be interpreted as category-level representation rather than mutually exclusive proportions. As summarised in Table 1, most documents were classified within Orthopedics and Sports Medicine (n = 16; 80%; 9 journals), predominantly published in Knee Surgery Sports Traumatology Arthroscopy, American Journal of Sports Medicine and Orthopaedic Journal of Sports Medicine. Sports Science (n = 11; 55%; 6 journals) was also highly represented, with Knee Surgery Sports Traumatology Arthroscopy, American Journal of Sports Medicine and Scandinavian Journal of Medicine & Science in Sports as the leading outlets. Medicine (miscellaneous) accounted for 10 documents (50%; 8 journals), mainly in American Journal of Sports Medicine, BMJ Open and Journal of Clinical Medicine, whereas Surgery (n = 7; 35%; 4 journals) was chiefly represented by Knee Surgery Sports Traumatology Arthroscopy, Clinical Orthopaedics and Related Research and Journal of Knee Surgery.

Rehabilitation-related fields such as *Physical Therapy* and *Sports Therapy and Rehabilitation* contributed 6 documents each (30%), with a stronger presence in American Journal of Sports Medicine, British Journal of Sports Medicine, Journal of Orthopaedic & Sports Physical Therapy and Scandinavian Journal of Medicine & Science in Sports. Less frequent categories included *Pharmacology (medical)* (Trials; n = 1; 5%) and pediatric-related areas (*Pediatrics*, *Perinatology and Child Health*; both represented by Journal of Pediatric Orthopaedics, n = 1; 5% each). As categories are not mutually exclusive, individual journals and articles were often indexed in multiple subject areas.

Table 1. Summary of subject categories (SCImago) of documents and journals

Category	Documents	Percent	N journals	Main journals
Orthopedics and Sports Medicine	16	80	9	Knee Surgery Sports Traumatology Arthroscopy; American Journal of Sports Medicine; Orthopaedic Journal of Sports Medicine
Sports Science	11	55	6	Knee Surgery Sports Traumatology Arthroscopy; Scandinavian Journal of Medicine & Science in Sports; American Journal of Sports Medicine
Medicine (miscellaneous)	10	50	8	BMJ Open; Journal of Clinical Medicine; Trials
Surgery	7	35	4	Knee Surgery Sports Traumatology Arthroscopy; Journal of Knee Surgery;

				Journal of Orthopaedic Surgery and Research
Physical Therapy	6	30	4	Journal of Orthopaedic & Sports Physical Therapy; British Journal of Sports Medicine; Scandinavian Journal of Medicine & Science in Sports
Sports Therapy and Rehabilitation	6	30	6	Journal of Orthopaedic & Sports Physical Therapy; British Journal of Sports Medicine; Scandinavian Journal of Medicine & Science in Sports
Pharmacology (medical)	1	5	1	Trials
Pediatrics	1	5	1	Journal of Pediatric Orthopaedics
Perinatology and Child Health	1	5	1	Journal of Pediatric Orthopaedics

3.3. SCIENTIFIC JOURNALS

The 20 documents included in this review were published across 13 different journals (Supplementary Table 3). The most prolific outlets were Knee Surg Sports Traumatol Arthrosc (n = 4; 20%), American Journal of Sports Medicine (n = 3; 15%) and the Orthopaedic Journal of Sports Medicine (n = 3; 15%), while all remaining journals contributed a single document each (5% per journal).

In terms of impact indicators, most journals were indexed in the first quartile (Q1, n = 11), with only two journals classified as Q2 (*Journal of Pediatric Orthopaedics* and *Trials*). The SJR values ranged from 0.655 to 4.724, indicating that the majority of the publications were disseminated in high-ranked outlets within orthopedics, sports medicine and sports science. When weighting the 2024 Journal Impact Factor by the number of documents per journal, the corpus yielded a weighted mean IF of 4.22, suggesting that research on this topic is predominantly published in high-impact journals rather than being dispersed in lower-impact sources.

3.4. Most cited documents

As shown in Figure 3, the distribution of total citations by year of publication shows a clear peak in 2012, when a single article accounted for 160 citations, markedly higher than in any other year. Moderate citation counts are observed for studies published in 2002 and 2017–2019 (ranging from 67 to 91 citations), indicating that work from this period has exerted sustained influence on the field. In contrast, documents published from 2021 onwards exhibit substantially fewer citations, with no citations yet recorded for articles from 2024 and 2025, which is consistent with the shorter time window for these more recent publications to be cited.

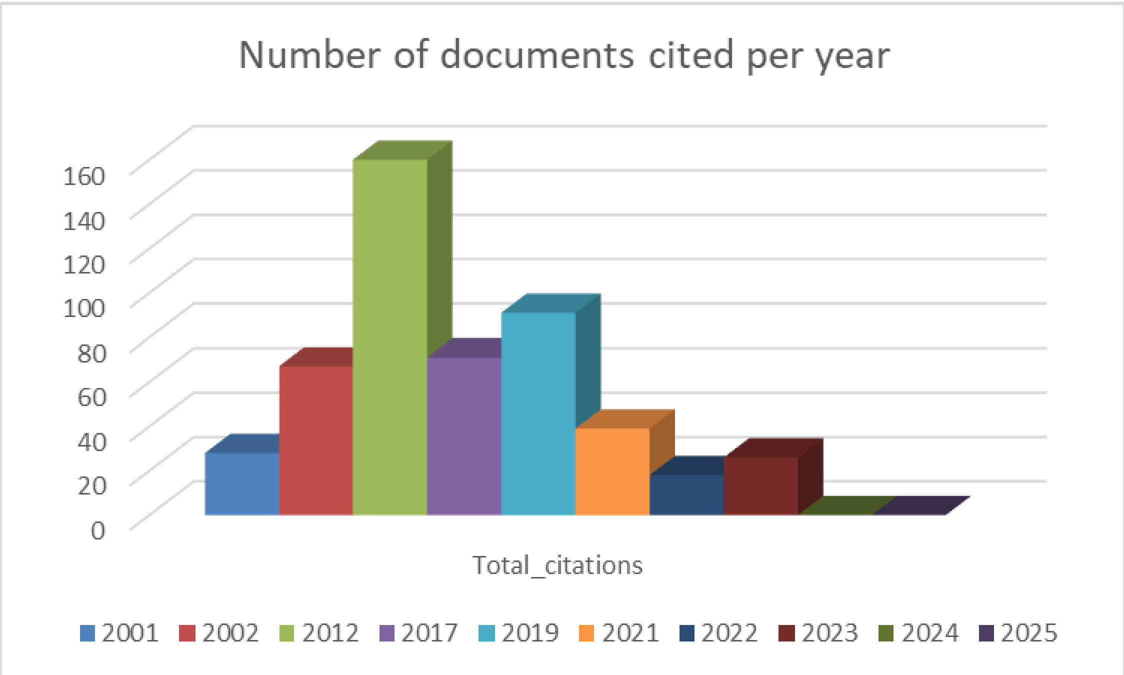


Figure 3. Total citations by year of publication

The citation distribution across outcome measures is clearly skewed towards a small number of instruments (as shown on Supplementary Figure 1). The IKDC-SKF (Subjective) shows the highest citation count (n=160), followed by the Lysholm score and the KOS-SAS, which also present relatively high citation frequencies. Intermediate levels of citation are observed for the KOS-ADLS, ACL-RSI, Kujala (AKPS) and Tegner Activity Scale, whereas the ACL-QOL, Pedi-IKDC and KOOS (global) have more modest citation numbers. The Cincinnati Knee Rating System is the least cited measure. Overall, these data indicate that research on knee-related outcomes has predominantly focused on a small set of subjective knee function scores, particularly the IKDC-SKF and Lysholm scales.

3.5. AUTHORS AND CO-AUTHORSHIP NETWORK

A total of 123 unique authors were identified across the included studies. Only a small proportion of these researchers contributed to more than one document, while the vast majority appeared as authors of a single paper. The co-authorship network displayed in Supplementary Figure 2 reveals several small clusters with limited interconnections, indicating that collaborations tend to be confined to relatively isolated research groups rather than involving a strongly integrated author community. When the temporal dimension is considered, earlier publications are mainly located in more peripheral clusters, while the densest groups of collaborations are coloured toward the more recent years, indicating that multi-author networks have become more active and interconnected in the most recent period of the literature.

3.6. COUNTRY-LEVEL PUBLICATION OUTPUT

Considering the country of the first author, 11 different countries contributed to the 20 documents included in this review (Table 2). When country output was operationalized using the first author's country (n = 20 publications), 11 countries were represented. The United States accounted for the highest share of first-author output (6/20 publications), followed by Australia (3/20), while France and Norway each contributed two first-author publications (2/20). The remaining countries (Canada, China, Germany, India, Singapore, Turkey, and the United Kingdom) each contributed a single first-author publication (1/20). This distribution is visually reflected in the VOSviewer density map, where the United States and Australia show the highest concentration of country-level publication density (Figure 4). In this first-author approach, total link strength (TLS)—defined as the cumulative strength of co-authorship links between countries in the VOSviewer network—is not expected to be informative (TLS = 0 across countries), because the metric reflects co-authorship links based on multi-country affiliations rather than single-country attribution to the first author. Overall, these results indicate that the corpus is led primarily by the United States and Australia, with additional but smaller first-author contributions distributed across Europe and Asia-Pacific.

Table 2. Representation of documents published by country and respective citations

Country	**Documents**	**Total citations**
Australia	3	26
Canada	1	15
China	1	19
France	2	46
Germany	1	23
Great Britain	1	3
India	1	0
Norway	2	35
Singapore	1	18
Turkey	1	10
USA	6	305

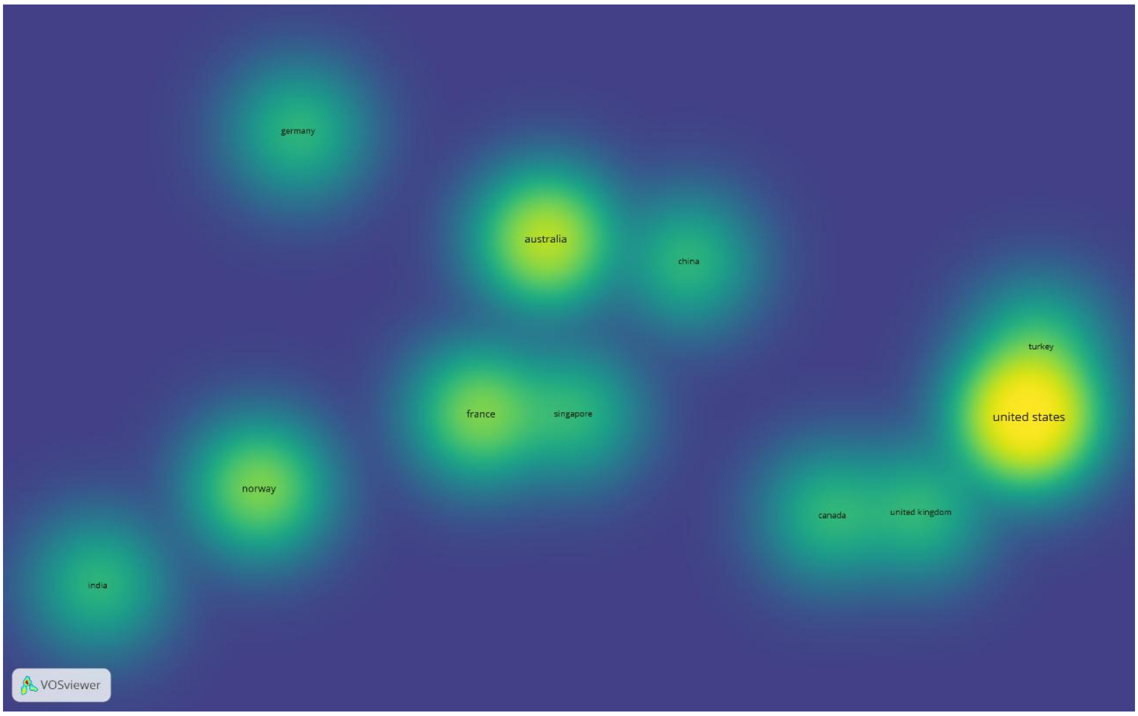


Figure 4. VOSviewer density visualisation of publications by country. Greater density indicates a higher number of publications.

Across the 20 included publications, the country-level co-authorship analysis based on all author affiliations identified 14 contributing countries (24 country–document attributions, reflecting multi-country papers). The network reveals a core collaborative structure dominated by the United States (7 documents; total link strength [TLS]=4) and Australia (3 documents; TLS=5), with Australia acting as a key connector to other collaborating countries. A tight European sub-network is evident, with Germany (2 documents; TLS=4), Denmark (1 document; TLS=4), and Sweden (1 document; TLS=4) showing comparatively strong link strengths despite lower output, indicating repeated international co-authorship ties within the corpus. In contrast, several countries contributed publications but exhibited no collaboration links within this dataset (TLS=0; e.g., Canada, China, France, Norway, Singapore, Turkey), suggesting predominantly single-country publications or collaborations that do not overlap with other countries represented in this corpus – Supplementary Figure 3. Additional peripheral contributions with minimal connectivity were observed (e.g., New Zealand, Saudi Arabia, and the United Arab Emirates; TLS=1), reinforcing an overall pattern of localized production with a small, internationally connected core.

3.7. ORGANISATION-LEVEL PUBLICATION OUTPUT

The organisation-level co-authorship mapping based on author affiliations – Supplementary Figure 4 - indicates a highly fragmented institutional structure, with all identified organisations contributing a single publication (Documents = 1) and no citation information recorded in the imported dataset (Citations = 0). Connectivity is therefore driven primarily by multi-affiliation papers, as reflected by the Total Link Strength (TLS) distribution. The strongest and most cohesive component is a dense Australian network (TLS = 11), linking multiple academic and clinical entities (e.g., University of Melbourne and University of Sydney–associated units, together with specialist sports medicine centres, hospitals, physiotherapy clinics and diagnostic imaging services), consistent with a multi-site, multi-institution collaboration concentrated within a single study. Additional mid-level components are observed with TLS = 6 and TLS = 5, representing separate institutional groupings that likely correspond to individual multi-institution publications, while smaller collaborative sets are captured at TLS = 4 and TLS = 3 (e.g., a coherent Norwegian grouping including Haukeland University Hospital/University of Bergen-related institutions at TLS = 3). Finally, a number of organisations display minimal connectivity (TLS = 1–2), indicating limited overlap of affiliations across the dataset. Overall, the organisation network suggests that the literature is produced by multiple independent institutional teams, with connectivity largely arising from a small number of highly multi-affiliated studies rather than repeated inter-institution collaboration across publications.

3.8. AUTHOR KEYWORDS

As shown in Figure 5, the keyword co-occurrence network of author keywords (minimum occurrence = 1) reveals a markedly skewed thematic structure dominated by an ACL–return-to-sport (RTS) core. In the summary table, only two keywords reach the highest frequency ("return to sport" and "anterior cruciate ligament," 5 occurrences each), followed by a small set of recurrent terms ("knee," "anterior cruciate ligament reconstruction," and "rehabilitation," 3 occurrences), whereas most remaining keywords appear only once (total link strength = 1). This distribution indicates that, within the included athlete-focused literature, perceived knee instability is primarily conceptualised through ACL injury/reconstruction pathways and their functional sequelae, reflected by closely related terms spanning clinical management and recovery (e.g., rehabilitation, physical therapy, autograft, anterolateral rotary instability, pivot shift, joint laxity, knee stability, and functional recovery). In contrast, alternative instability phenotypes—such as patellofemoral instability—are represented only peripherally, with “patellar instability,” “lateral patellar dislocation,” and “MPFL/MPFL reconstruction” each occurring once, suggesting limited thematic centrality in this corpus. Finally, minor synonym/format fragmentation (e.g., “orthopaedics” vs “orthopedics,” and multiple ACL/ACLR variants) may dilute individual keyword frequencies; nevertheless, the overall network remains clearly anchored around ACL/ACLR, rehabilitation, and RTS as the dominant conceptual axis.

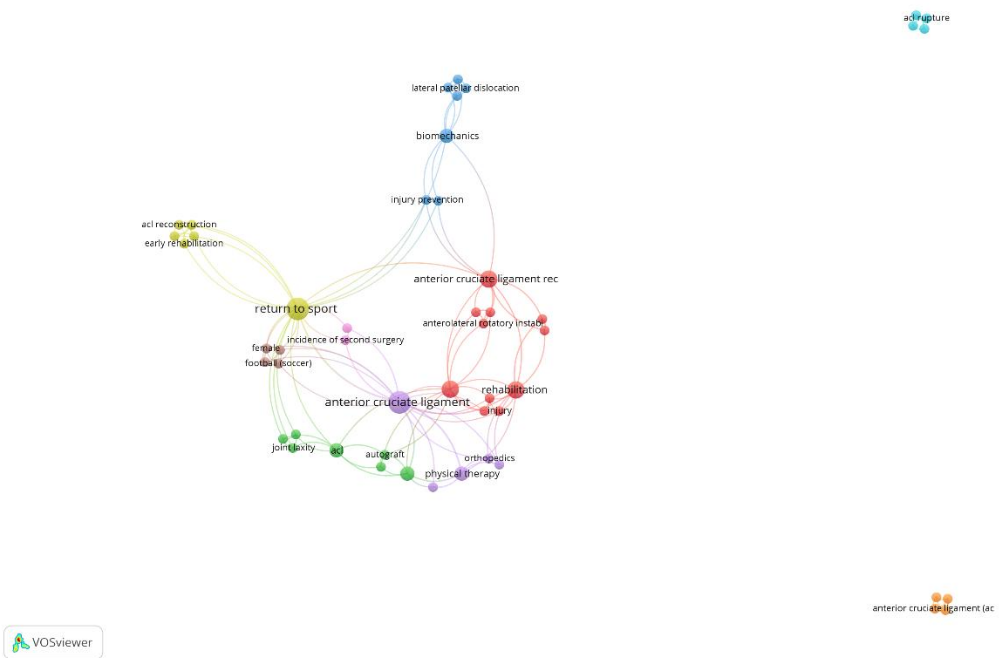


Figure 5. Keyword co-occurrence network using a minimum threshold of two occurrences

3.9. PSYCHOMETRIC SYNTHESIS

In this review, each instrument was appraised across three complementary domains: (i) structural and descriptive characteristics of the questionnaire, including item structure, scoring procedures, recall period, and intended population; (ii) a concise synthesis of psychometric evidence; and (iii) an integrated interpretive summary of its relevance for athlete-focused knee instability assessment. The structural characteristics of the included PROMs are summarised in Table 3, whereas the main psychometric findings are synthesised in Table 4. Given the breadth and heterogeneity of the available validation literature, detailed PROM-by-PROM descriptions, study-level psychometric extractions, and COSMIN-based appraisals are reported in Supplementary Tables S4–S5, thereby preserving transparency while allowing a more concise and interpretable synthesis in the main manuscript.

Internal consistency was extracted primarily using Cronbach's alpha (α), adopting 0.70 as the minimum acceptable threshold for research purposes (Nunnally, 1978). However, in accordance with COSMIN principles, alpha coefficients were interpreted cautiously when evidence of unidimensionality or structural validity was absent. Therefore, internal consistency values were retained descriptively in Supplementary Table S5, but were not automatically interpreted as favourable evidence in the summary synthesis. Indices of temporal stability, measurement error, and model fit were reported when available. Test–retest correlations and intraclass correlation coefficients (ICCs) were interpreted according to Vincent and Weir, with values >0.90 considered high, 0.80–0.90 moderate, and <0.80 inadequate (Vincent & Weir, 2012). This approach allowed psychometric findings to be interpreted consistently despite heterogeneity in populations, language versions, and validation designs.

Overall, the included PROMs did not perform equivalently across constructs or evaluative purposes. The strongest and most consistent favourable evidence was observed for reliability, whereas favourable evidence for content validity, measurement error, structural validity, cross-cultural validity, and responsiveness was more limited or absent for several instruments. Broad knee instruments such as the IKDC-SKF, KOOS, and, historically, Lysholm, showed the most established role as symptom- and function-oriented measures, with reproducibility evidence reported across several validation contexts (Larsen et al., 2023; Lysholm & Gillquist, 1982; Tigerstrand Grevnerts et al., 2017; Wang et al., 2016a). However, this evidence should be interpreted cautiously because favourable reliability does not necessarily imply comprehensive measurement adequacy. In particular, athletic populations may present ceiling effects, high baseline function, and sport-specific demands that are not fully captured by broad function scores. This is especially relevant for legacy tools such as Lysholm and Tegner, which remain widely used but may require cautious interpretation when subtle change in high-functioning athletes is of interest (K. K. Briggs et al., 2009; Tegner & Lysholm, 1985; Wang et al., 2016b).

By contrast, more domain-specific instruments capture constructs that are not fully represented in broad function scales. The ACL-RSI was specifically developed to assess psychological readiness to return to sport and showed favourable evidence for hypotheses testing and reliability in translated versions, although formal cross-cultural validity evidence was not demonstrated in the current synthesis (Albano et al., 2022; Piussi et al., 2025; Webster & Feller, 2018; Webster et al., 2008). The ACL-QOL provides complementary information on the broader quality-of-life consequences of ACL injury and reconstruction, but the favourable COSMIN-based evidence remains uneven across properties, with stronger support for selected aspects of validity and reliability than for measurement error, structural validity, or cross-cultural validity (Kinikli et al., 2015; Mohtadi, 1998; Silva et al., 2018). The Tegner Activity Scale remains useful for contextualising sport exposure and participation level rather than for quantifying symptoms or functional deficit per se (K. K. Briggs et al., 2009; Tegner & Lysholm, 1985).

Taken together, the psychometric synthesis supports a concise core-set logic, in which PROM selection is guided by construct relevance and intended application rather than by historical usage alone. Within this framework, broad symptom/function instruments are best interpreted as anchors, whereas psychological readiness, activity exposure, and broader quality-of-life consequences may require complementary assessment depending on the aims of the study. This perspective is consistent with contemporary COSMIN principles, which emphasise that PROM selection should be based primarily on content validity and fitness for purpose rather than frequency of use alone (Mokkink et al., 2024; Terwee et al., 2018). This psychometric synthesis informed the integrative interpretation presented in Section 4.6, in which bibliometric prominence and measurement performance were considered jointly to identify the most defensible PROM options for athlete-focused knee instability assessment.

Table 3. Structural Characteristics of Included PROMs

PROM	Year	Developer	Items	Domains / subscales	Response format	Score range	Recall period	Typical population

ACL-QOL	1998	Mohtadi (Mohtadi, 1998)	32	Symptoms/physical complaints; work-related concerns; recreational activities and sport; lifestyle modification; social/emotional impact	100-mm visual analogue scale	0–100	Past week	ACL-deficient or ACL-reconstructed patients
ACL-RSI	2008	Webster et al. (Webster et al., 2008)	12	Emotions; confidence in performance; risk appraisal	Likert-type items	0–100	Current return-to-sport readiness	Patients after ACL reconstruction
CKRS	1989	Noyes et al. (Noyes et al., 1989)	19	Symptoms; functional limitations; sport activity; occupational activity	Item-specific scoring	Composite score	Variable / not consistently specified	Ligament injury populations
IKDC-SKF	2001	Irrgang et al. (Irrgang et al., 2001)	18	Symptoms; sports activity; knee function	Item-dependent ordinal responses	0–100	Past 4 weeks	Broad knee injury population
KOOS	1998	Roos et al. (Roos et al., 1998)	42	Pain; symptoms; activities of daily living; sport/recreation; knee-related quality of life	Likert scale (0–4)	0–100 per subscale	Past week	Knee injury and early osteoarthritis populations
AKPS / Kujala	1993	Kujala et al. (Kujala et al., 1993)	13	Pain and functional limitations related to patellofemoral disorders	Weighted ordinal items	0–100	Current status	Patellofemoral pain / patellofemoral disorder populations
Pedi-IKDC	2011	Kocher et al. (Kocher et al., 2011)	19	Symptoms; sports activity; knee function	Item-dependent ordinal responses	0–100	Past 4 weeks	Paediatric and adolescent knee injury populations
Tegner Activity Scale	1985	Tegner & Lysholm (Tegner & Lysholm, 1985)	1	Activity level	Ordinal activity rating scale	0–10	Current activity level	Knee ligament injury populations
Lysholm Knee Score	1982	Lysholm & Gillquist (Lysholm & Gillquist, 1982)	8	Limp; support; locking; instability; pain; swelling; stair climbing; squatting	Weighted ordinal scoring	0–100	Current status	Ligament injury populations

Note 1. Abbreviations: ACL, anterior cruciate ligament; AKPS, Anterior Knee Pain Scale; CKRS, Cincinnati Knee Rating System; IKDC-SKF, International Knee Documentation Committee Subjective Knee Form; KOOS, Knee injury and Osteoarthritis Outcome Score; PROMs, patient-reported outcome measures. Only the nine core PROMs retained in the main psychometric synthesis are included in the main manuscript. Derivative, domain-specific, or supplementary forms (e.g., KOS-ADLS, KOS-SAS, and isolated KOOS subscales) are not treated as independent core instruments in the main synthesis.

Table 4. Summary of Psychometric Evidence Across PROMs

PROM	Validation studies (n)	Content validity	Internal consistency	Reliability	Measurement error	Structural validity	Hypotheses testing	Cross-cultural adaptation	Cross-cultural validity	Responsiveness
ACL-QOL	4	1/4	0/4	2/4	0/4	0/4	2/4	1/4	0/4	1/4
ACL-RSI	3	0/3	0/3	2/3	0/3	0/3	3/3	2/3	0/3	0/3
CKRS	2	1/2	0/2	2/2	0/2	0/2	1/2	0/2	0/2	1/2
IKDC-SKF	7*	1/7	0/7	6/7	0/7	0/7	3/7	3/7	0/7	1/7
KOOS	5	0/5	0/5	4/5	0/5	0/5	2/5	1/5	0/5	1/5
AKPS / Kujala	6	0/6	0/6	5/6	0/6	2/6	5/6	5/6	0/6	2/6
Pedi-IKDC	3	1/3	0/3	3/3	0/3	0/3	1/3	3/3	0/3	1/3
Tegner Activity Scale	4	0/4	0/4	4/4	0/4	0/4	4/4	2/4	0/4	2/4
Lysholm Knee Score	3	0/3	0/3	2/3	0/3	0/3	1/3	0/3	0/3	1/3

Note 2. Denominators represent the total number of validation studies listed for each PROM in that sheet. Cross-cultural validity was not supported by favourable evidence in the current extraction matrix for any of the nine core PROMs and is therefore shown as 0/n across instruments. *Letchford et al. (2015) was retained under IKDC-related evidence because the study evaluated the IKDC activity/participation item as a distinct participation measure. However, these findings refer only to this single IKDC item and should not be interpreted as evidence for the full IKDC-SKF total score.

4. DISCUSSION

The present review sought to interpret the bibliometric profile of questionnaires used to assess functional knee instability in sporting contexts and to critically appraise the most frequently cited instruments using their reported psychometric evidence. Across the 20 eligible documents, we observed (i) a clear post-2017 acceleration of output, (ii) a publication ecosystem concentrated in high-ranked orthopaedics/sports medicine outlets, (iii) a highly skewed citation structure favouring a small set of legacy, function-oriented PROMs (particularly IKDC-SKF and Lysholm) and, (iv) collaboration patterns that are fragmented at author and institutional levels, with connectivity driven by a limited number of multi-affiliation studies. A key interpretive point emerging from this combined bibliometric–psychometric approach is that "bibliometric prominence" is not synonymous with "measurement optimality." In measurement science, the COSMIN framework positions content validity (relevance, comprehensiveness, comprehensibility) as foundational; without it, even highly reliable tools can systematically miss the intended construct (Mokkink et al., 2024).

Therefore, the core contribution of this paper is not simply identifying the most *visible* instruments, but clarifying where instrument visibility aligns, or fails to align, with measurement adequacy for athlete-centred instability assessment and return-to-sport decision-making.

4.1. FIELD MATURATION, JOURNAL ECOLOGY, AND WHY LEGACY PROMS DOMINATE

Annual publication trends show an early sparse phase (2001–2012) followed by a pronounced increase from 2017 onwards, with 17 of 20 publications occurring between 2017 and 2025 and peaks of four publications in 2017 and 2019. This pattern likely reflects the broader consolidation of outcome-based practice in sports knee care, where longitudinal monitoring, return-to-sport (RTS) decision-making, and comparative effectiveness research require standardized PROM deployment across rehabilitation milestones. The field's topical emphasis on RTS is further supported by keyword mapping, in which "return to sport" and "treatment outcome" form recurrent, central constructs once generic indexing terms are accounted for. This matches the wider clinical literature in which RTS after ACL injury is treated as a complex, multidimensional endpoint that requires integration of symptoms/function, psychological readiness, and contextual factors, frequently formalised through consensus processes (Meredith et al., 2020).

Importantly, the temporal profile should also be interpreted in light of broader citation dynamics. Recent publications, particularly those from 2024–2025, predictably show low or absent citation counts because they have had limited time to circulate within the scientific community. However, citation-based prominence is not determined by time alone and may also reflect cumulative advantage mechanisms, commonly described as the Matthew effect in science, whereby earlier, highly visible or already influential works tend to accumulate disproportionate citation advantages over time. Therefore, the dominance of legacy PROMs in the present bibliometric profile should be interpreted as evidence of scientific diffusion, historical visibility and field uptake, rather than as direct evidence of superior psychometric quality. Conversely, newer instruments or recent validation studies may be underrepresented in citation-based analyses despite their potential methodological relevance. In line with responsible-metrics frameworks, citation indicators should therefore be used to contextualise influence and dissemination, but not to infer the intrinsic quality, clinical utility or measurement validity of individual studies (Larivière et al., 2016; Wilsdon, 2015).

4.2. DISCIPLINARY CONCENTRATION AND JOURNAL ECOSYSTEM: WHY ORTHOPAEDICS/SPORTS MEDICINE DOMINATE

Most documents were indexed within Orthopedics and Sports Medicine (80%) and Sports Science (55%), with meaningful representation across rehabilitation-related categories (e.g., Physical Therapy; Sports Therapy and Rehabilitation; 30% each). The journal profile mirrors this disciplinary anchoring: 20 studies were distributed across 13 journals, led by Knee Surgery Sports Traumatology Arthroscopy (20%), American Journal of Sports Medicine (15%), and Orthopaedic Journal of Sports Medicine (15%), and predominantly Q1-ranked outlets with a weighted mean impact factor of 4.22. Collectively, this supports the interpretation that sport-related instability outcomes are largely framed within the orthopaedic–sports medicine continuum (injury mechanisms, reconstruction, rehabilitation, RTS), rather than being primarily driven by instrument-development communities. This publication environment tends to privilege instruments that are feasible, widely recognized, and comparable across cohorts—conditions that historically favour legacy PROMs such as IKDC and Lysholm.

This framing is also consistent with the epidemiologic reality that ACL injury represents a high-burden condition in sports populations, particularly in contact and limited-contact sports with substantial athlete exposure and sustained research infrastructure. A large systematic review and meta-analysis reports ACL injury incidence rates in contact sports and highlights sex-based differences, reinforcing the prominence of ACL-related cohorts as a driver of outcomes research volume and visibility (Montalvo et al., 2019). In practice, the combination of high incidence, high functional impact, and a relatively standardised clinical pathway (surgery → rehabilitation → RTS evaluation) creates repeated measurement timepoints that naturally promote PROM usage and reporting.

4.3. Citation structure and instrument "core dominance": bibliometric prominence versus construct coverage

The most cited-document analysis revealed a strong citation peak in 2012, with a single article accounting for 160 citations. At the instrument level, citation distribution is clearly skewed toward a small set of outcome measures: IKDC-SKF has the highest citation count (=160), followed by the Lysholm score and KOS-SAS; intermediate prominence is observed for KOS-ADLS, ACL-RSI, Kujala (AKPS), and Tegner, whereas ACL-QOL, Pedi-IKDC, and KOOS (global) are less prominent, and the Cincinnati Knee Rating System is least cited.

A plausible interpretation is that IKDC-SKF and Lysholm dominate because they function as broadly applicable, historically entrenched "default endpoints" for knee outcomes in sports medicine. However, this dominance should be interpreted as evidence of historical uptake and comparability across cohorts rather than as direct evidence of superior measurement quality. The psychometric synthesis shows that favourable evidence is not uniformly distributed across measurement properties: reliability is comparatively well represented, whereas content validity, measurement error, cross-cultural validity, and responsiveness remain incompletely supported for several instruments.

However, dominance by legacy tools raises two measurement concerns. First, legacy function scores can show ceiling effects in higher-functioning athletic populations, which constrains responsiveness and threatens interpretability when the aim is to discriminate near full recovery. Evidence comparing IKDC and Lysholm has highlighted ceiling-effect concerns in outcome assessment after ACL reconstruction, particularly for the Lysholm scale (Ra et al., 2014).

Second, bibliometric dominance does not guarantee superiority in change detection. This interpretation is reinforced by external evidence that more recently developed ACL-specific PROMs may outperform legacy tools in responsiveness. For example, Comins et al. reported that KNEES-ACL demonstrated superior responsiveness compared to commonly used measures such as KOOS, IKDC, and Lysholm in ACL reconstruction cohorts, leveraging modern psychometric development approaches (including Rasch/IRT methods) (Comins et al., 2018).

This misalignment is also consistent with broader measurement-science guidance that instrument selection should be anchored in content validity and intended use (discriminative vs evaluative; individual vs group decision-making), particularly when publication ecosystems favour established measures for comparability.

4.4. COLLABORATION NETWORKS (AUTHORS, ORGANISATIONS, COUNTRIES): IMPLICATIONS FOR MEASUREMENT CONSISTENCY

The co-authorship analysis identified 123 unique authors, with the majority contributing to a single paper and forming small, weakly interconnected clusters. Temporally, more recent studies tend to concentrate in denser collaboration clusters, suggesting increased network activity in the recent period. At country level (first-author attribution), 11 countries contributed to the 20 documents, led by the United States (6/20; 305 citations) and Australia (3/20), with smaller contributions across Europe and Asia-Pacific. When using all affiliations (co-authorship by countries), a small internationally connected core emerges, dominated by the United States and Australia, with a European sub-network showing relatively strong link strengths despite lower output.

At organisational level, the network is highly fragmented, with connectivity largely driven by a small number of multi-affiliation papers; the strongest component is a dense Australian network (TLS = 11), consistent with a single multi-site collaboration rather than repeated inter-institution partnerships across studies. Methodologically, this fragmentation can contribute to heterogeneity in PROM selection, administration timing, language versions, and reporting depth of interpretability metrics (e.g., SEM/SDC/MDC). As a result, the literature may accumulate instrument usage without consistently accumulating comparable psychometric evidence, which complicates the goal of naming a single "best" PROM based on pooled properties.

4.5. Keyword maps and indexing effects: what the literature "talks about" versus what it "measures"

The resulting co-occurrence structure is markedly skewed toward an ACL–return-to-sport (RTS) core, with only two terms reaching the highest frequency ("return to sport" and "anterior cruciate ligament", 5 occurrences each), followed by a limited second tier ("knee", "anterior cruciate ligament reconstruction", and "rehabilitation", 3 occurrences), while most remaining keywords occur only once (TLS = 1). This distribution indicates a corpus in which thematic convergence exists mainly around a small conceptual axis (ACL/ACLR → rehabilitation → RTS), whereas peripheral topics are represented by sparse, low-frequency terminology—an expected pattern in small corpus and in author-keyword layers where synonymy and idiosyncratic keywording practices can dilute co-occurrence signals.

Conceptually, the ACL–RTS dominance is consistent with sport-exposure epidemiology and the clinical research infrastructure surrounding ACL injury. Recent umbrella evidence synthesising sport-specific meta-analyses continues to show that sport-related ACL injuries have a high incidence in several high-exposure sports (e.g., football/soccer, basketball, volleyball), which sustains large cohorts and repeated outcome measurement timepoints, thereby amplifying ACL-centred outcome research (Martinez-Calderon et al., 2025). In the present corpus, "instability" is therefore predominantly operationalised through ACL injury/reconstruction trajectories and their functional sequelae (e.g., pivot shift, anterolateral rotary instability, joint laxity, knee stability, functional recovery), with relatively limited thematic centrality for alternative instability phenotypes. This is not a claim that patellofemoral instability is clinically irrelevant in athletes—indeed, epidemiologic data in athletic cohorts show measurable injury rates and sports-related occurrence for patellar instability/dislocation—but rather that such pathways are underrepresented in this athlete-focused PROM corpus compared with the ACL literature (Mitchell et al., 2015; Raney et al., 2025).

The prominence of “return to sport” within the author-keyword map aligns with consensus work that frames RTS not as a single clearance decision but as a continuum embedded within rehabilitation and risk management (Ardern et al., 2016; Meredith et al., 2020). This has direct implications for PROM interpretation: if RTS is a dominant organising concept of the literature, then PROM selection should reflect multi-domain readiness (symptoms/function, psychological readiness, participation context), rather than relying solely on historically entrenched function scores. This reinforces the methodological rationale of Section 3.9 approach (psychometric appraisal and construct alignment), and is consistent with PROM evaluation guidance that prioritises content validity and construct coverage over frequency of use (Terwee et al., 2018).

4.6. Integrating psychometrics with bibliometrics: how we interpret “best” PROMs for athlete instability

The psychometric synthesis presented in Section 3.9 indicates that several knee PROMs demonstrate favourable reliability evidence across populations and language versions; however, they differ substantially in construct alignment, interpretability, and completeness of evidence for athlete-focused instability assessment. In this context, IKDC-SKF appears to be the most defensible broad symptom/function anchor because of its historical uptake and comparatively consistent reliability evidence. Lysholm, by contrast, should be interpreted primarily as a historically entrenched legacy score: although familiar and widely used, its weighted item structure and documented ceiling effects may reduce sensitivity when the objective is to detect subtle late-stage recovery or near-return-to-performance change in athletic populations, where residual deficits may persist despite high total scores (Ra et al., 2014).

Because return to sport is a central organizing outcome in this corpus, psychological readiness should be treated as a necessary complementary measurement domain when RTS decision-making is a primary endpoint. Consistent with the Panther Symposium consensus, RTS is multidimensional and cannot be captured by function scores alone, supporting a complementary pairing strategy—e.g., a function/symptom PROM such as IKDC-SKF combined with an RTS-related psychological measure such as ACL-RSI—rather than reliance on a single instrument to operationalise “functional instability in sport” (Diermeier et al., 2021). Under this interpretation, symptom- and function-oriented PROMs provide necessary information but remain insufficient in isolation when the assessment target includes confidence, perceived risk, sport exposure, and readiness for high-demand tasks.

From a measurement-selection perspective, bibliometric prominence should not be equated with psychometric superiority. COSMIN guidance emphasizes that PROM recommendation must be primarily grounded in content validity and supported by adequate evidence regarding reliability, measurement error, and responsiveness (Mokkink et al., 2024). Within this framework, the “best PROM” is not simply the most frequently used instrument but rather the instrument—or minimal set of instruments—that most appropriately captures the intended construct within the relevant population and evaluative context. This distinction is particularly relevant in athlete-focused knee instability research, where the construct extends beyond symptoms alone and encompasses performance capacity, psychological readiness, and participation-level outcomes.

This perspective also helps reconcile the dominance of legacy tools in the bibliometric record with emerging evidence suggesting that newer ACL-specific instruments may demonstrate superior responsiveness in certain contexts. For example, responsiveness analyses have indicated that KNEES-ACL may detect meaningful clinical change more effectively than several commonly used knee PROMs (Comins et al., 2018). This suggests that the instruments most frequently represented in the bibliometric core may not necessarily correspond to the most refined solutions for measurement optimisation, particularly when the assessment target is late-stage, athlete-centred instability rather than general knee function.

Beyond responsiveness alone, this observation raises a broader construct-level issue. Functional knee instability in sport should not be conceptualized as a unidimensional symptom or as the mere subjective correlate of ligament laxity. In athlete-centered contexts, the construct is inherently multidimensional, combining perceived episodes of giving way or loss of control, confidence in the knee, sport-specific movement tolerance, psychological readiness, and the ability to perform high-demand tasks such as cutting, pivoting, deceleration, landing, acceleration, and contact- or opponent-constrained actions. Therefore, the measurement requirements for sport-specific application are likely to differ from those of general knee outcome assessment. Instruments intended for athletic populations should demonstrate sufficient content validity for late-stage rehabilitation and return-to-sport contexts, including items that are sensitive to near-ceiling performance, sport-specific instability triggers, multidirectional and high-velocity tasks, and the athlete's perceived capacity to reproduce these actions under fatigue, uncertainty, and competitive pressure. From this perspective, future PROM development should move beyond adapting broad knee function scales and should explicitly model the instability construct around the demands of sport participation. Such development may require item generation with athletes, clinicians, and sport-specific experts, followed by modern psychometric testing to confirm dimensionality, responsiveness, and interpretability in high-functioning athletic cohorts.

In the absence of PROMs specifically developed around sport-task instability, current instrument selection should be understood as a pragmatic approximation rather than a definitive measurement solution. A defensible interim strategy is to combine complementary instruments according to the intended purpose of assessment: IKDC-SKF as a broad symptom/function anchor, ACL-RSI when return-to-sport decision-making is central, Tegner when activity exposure needs to be contextualized, and ACL-QOL when broader ACL-related quality-of-life impact is a priority. Nevertheless, this approach remains compensatory, as it does not replace the need for instruments explicitly designed to capture the multidimensional and task-specific nature of functional knee instability in athletic populations.

4.7. Limitations and future directions

This review is constrained by the small bibliometric corpus (n=20), which can exaggerate the apparent dominance of a few instruments and collaboration clusters. In addition, citation accrual is influenced by exposure time and by broader citation dynamics, including cumulative advantage and field visibility effects (Merton, 1968; Wang, 2014). Therefore, low or absent citations for the most recent publications may reflect limited exposure time rather than limited scientific relevance, potentially underestimating emerging PROMs and newer psychometric evidence. Conversely, high citation counts for older PROMs may partly reflect historical diffusion, repeated uptake and the Matthew effect, rather than superior measurement quality. This limitation reinforces the need to interpret citation indicators as measures of visibility and dissemination, not as direct proxies for the methodological quality, clinical utility or validity of individual studies (Larivière et al., 2016; Wilsdon, 2015). This is particularly relevant in human movement research, where compressed evaluation cycles and metric-oriented academic cultures may distort how time, impact and knowledge quality are interpreted (Ramalho et al., 2025). Second, the collaboration maps indicate fragmented author and institutional structures with connectivity driven by a limited number of multi-affiliation papers; this fragmentation can propagate heterogeneity in PROM choice, administration timing, and completeness of interpretability reporting (e.g., SEM/SDC/MDC), which constrains robust cross-study comparisons of "best" PROM performance.

Future research should prioritise (i) consistent reporting of measurement error, including SEM, SDC/MDC and their interpretation against MIC or MCID thresholds; (ii) explicit justification of PROM selection based on content validity for sport-related instability constructs; (iii) formal cross-cultural validity testing using appropriate approaches such as differential item functioning, measurement invariance, IRT, Rasch analysis, or multi-group confirmatory factor analysis; (iv) explicit mapping of PROM domains to RTS outcomes; and (v) development of sport-specific instability instruments capable of capturing multidirectional, high-velocity, fatigue-sensitive, and psychologically demanding tasks typical of athletic participation. These steps would reduce heterogeneity and strengthen comparability across cohorts and sports, aligning PROM use with contemporary measurement standards and reporting expectations.

BIBLIOGRAPHY

Abulhasan, J. F., & Grey, M. J. (2017). Anatomy and Physiology of Knee Stability. *Journal of Functional Morphology and Kinesiology*, *2*(4), 34. https://doi.org/10.3390/jfmk2040034

Albano, T. R., Lima, P. O. P., Rodrigues, C. A. S., Melo, A. K. P., Tavares, M. L. A., & Almeida, G. P. L. (2022). Measurement properties of the Brazilian Portuguese anterior cruciate ligament - return to sport after injury (ACL-RSI) scale short version after anterior cruciate ligament reconstruction. *Braz J Phys Ther*, *26*(4), 100421. https://doi.org/10.1016/j.bjpt.2022.100421

Ardern, C. L., Glasgow, P., Schneiders, A., Witvrouw, E., Clarsen, B., Cools, A.,…Bizzini, M. (2016). 2016 Consensus statement on return to sport from the First World Congress in Sports Physical Therapy, Bern. *Br J Sports Med*, *50*(14), 853-864. https://doi.org/10.1136/bjsports-2016-096278

Barber-Westin, S. D., Noyes, F. R., & McCloskey, J. W. (1999). Rigorous statistical reliability, validity, and responsiveness testing of the cincinnati knee rating system in 350 subjects with uninjured, injured, or anterior cruciate ligament-reconstructed knees [Article]. *American Journal of Sports Medicine*, *27*(4), 402-416. https://doi.org/10.1177/03635465990270040201

Bausch, N., Eyre, E., Pearce, G., & Palmer, S. (2025). Knee-Related Patient-Reported Outcome Measures for Young Adults: A Scoping Meta-Review. *Musculoskeletal Care*, *23*(1), e70037. https://doi.org/10.1002/msc.70037

Bengtsson, J., Möllborg, J., & Werner, S. (1996). A study for testing the sensitivity and reliability of the Lysholm knee scoring scale. *Knee Surg Sports Traumatol Arthrosc*, *4*(1), 27-31. https://doi.org/10.1007/bf01565994

Briggs, K. K., Lysholm, J., Tegner, Y., Rodkey, W. G., Kocher, M. S., & Steadman, J. R. (2009). The Reliability, Validity, and Responsiveness of the Lysholm Score and Tegner Activity Scale for Anterior Cruciate Ligament Injuries of the Knee: 25 Years Later. *The American Journal of Sports Medicine*, *37*(5), 890-897. https://doi.org/10.1177/0363546508330143

Bryant, D., Norman, G., Stratford, P., Marx, R. G., Walter, S. D., & Guyatt, G. (2006). Patients undergoing knee surgery provided accurate ratings of preoperative quality of life and function 2 weeks after surgery. *J Clin Epidemiol*, *59*(9), 984-993. https://doi.org/10.1016/j.jclinepi.2006.01.013

Buckinx, F., Bornheim, S., Remy, G., Van Beveren, J., Reginster, J., Bruyère, O.,…Kaux, J. F. (2019). French translation and validation of the "Anterior Knee Pain Scale" (AKPS). *Disabil Rehabil*, *41*(9), 1089-1094. https://doi.org/10.1080/09638288.2017.1419288

Celik, D., Coşkunsu, D., & Kiliçoğlu, O. (2013). Translation and cultural adaptation of the Turkish Lysholm knee scale: ease of use, validity, and reliability. *Clin Orthop Relat Res*, *471*(8), 2602-2610. https://doi.org/10.1007/s11999-013-3046-z

Chaharmahali, L., Gandomi, F., Yalfani, A., & Fazaeli, A. (2021). The effect of self-reported knee instability on plantar pressure and postural sways in women with knee osteoarthritis. *Journal of Orthopaedic Surgery and Research*, *16*(1), 677. https://doi.org/10.1186/s13018-021-02823-6

Clancy, R. B., Herring, M. P., & Campbell, M. J. (2017). Motivation Measures in Sport: A Critical Review and Bibliometric Analysis. *Front Psychol*, *8*, 348. https://doi.org/10.3389/fpsyg.2017.00348

Collins, N. J., Misra, D., Felson, D. T., Crossley, K. M., & Roos, E. M. (2011). 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 [Article]. *Arthritis Care and Research*, *63*(SUPPL. 11), S208-S228. https://doi.org/10.1002/acr.20632

Comins, J., Brodersen, J., Krogsgaard, M., & Beyer, N. (2008). Rasch analysis of the Knee injury and Osteoarthritis Outcome Score (KOOS): a statistical re-evaluation. *Scand J Med Sci Sports*, *18*(3), 336-345. https://doi.org/10.1111/j.1600-0838.2007.00724.x

Comins, J. D., Siersma, V. D., Lind, M., Jakobsen, B. W., & Krogsgaard, M. R. (2018). KNEES-ACL has superior responsiveness compared to the most commonly used patient-reported outcome measures for anterior cruciate ligament injury. *Knee Surg Sports Traumatol Arthrosc*, *26*(8), 2438-2446. https://doi.org/10.1007/s00167-018-4961-z

Cross, M. (1996). Clinical terminology for describing knee instability. *SPORTS MEDICINE AND ARTHROSCOPY REVIEW*, *4*(4), 313-318.

da Cunha, R. A., Costa, L. O., Hespanhol Junior, L. C., Pires, R. S., Kujala, U. M., & Lopes, A. D. (2013). Translation, cross-cultural adaptation, and clinimetric testing of instruments used to assess patients with patellofemoral pain syndrome in the Brazilian population. *J Orthop Sports Phys Ther*, *43*(5), 332-339. https://doi.org/10.2519/jospt.2013.4228

da Silva-Júnior, F. B., Dibai-Filho, A. V., Barros, D. C. C., Dos Reis-Júnior, J. R., Gonçalves, M. B. S., Soares, A. R.,…Pires, F. O. (2024). Anterior Knee Pain Scale (AKPS): structural and criterion validity in Brazilian population with patellofemoral pain. *BMC Musculoskelet Disord*, *25*(1), 39. https://doi.org/10.1186/s12891-024-07164-z

Dammerer, D., Liebensteiner, M. C., Kujala, U. M., Emmanuel, K., Kopf, S., Dirisamer, F., & Giesinger, J. M. (2018). Validation of the German version of the Kujala score in patients with patellofemoral instability: a prospective multi-centre study. *Arch Orthop Trauma Surg*, *138*(4), 527-535. https://doi.org/10.1007/s00402-018-2881-5

de Loës, M., Dahlstedt, L. J., & Thomée, R. (2000). A 7-year study on risks and costs of knee injuries in male and female youth participants in 12 sports. *Scand J Med Sci Sports*, *10*(2), 90-97. https://doi.org/10.1034/j.1600-0838.2000.010002090.x

Diermeier, T. A., Rothrauff, B. B., Engebretsen, L., Lynch, A., Svantesson, E., Hamrin Senorski, E. A.,…Musahl, V. (2021). Treatment after ACL injury: Panther Symposium ACL Treatment Consensus Group. *Br J Sports Med*, *55*(1), 14-22. https://doi.org/10.1136/bjsports-2020-102200

Ebrahimzadeh, M. H., Makhmalbaf, H., Golhasani-Keshtan, F., Rabani, S., & Birjandinejad, A. (2015). The International Knee Documentation Committee (IKDC) Subjective Short Form: a validity and reliability study. *Knee Surg Sports Traumatol Arthrosc*, *23*(11), 3163-3167. https://doi.org/10.1007/s00167-014-3107-1

Ekstrand, J., Hägglund, M., & Waldén, M. (2011). Injury incidence and injury patterns in professional football: the UEFA injury study. *Br J Sports Med*, *45*(7), 553-558. https://doi.org/10.1136/bjsm.2009.060582

Huang, H., Zhang, S., Wang, Y., Tegner, Y., Jiang, Y., & Ao, Y. (2022). Reliability and Validity of a Chinese Version of the Lysholm Score and Tegner Activity Scale for Knee Arthroplasty. *J Rehabil Med*, *54*, jrm00317. https://doi.org/10.2340/jrm.v54.2304

Hughes, J. D., Rauer, T., Gibbs, C. M., & Musahl, V. (2019). Diagnosis and treatment of rotatory knee instability. *Journal of Experimental Orthopaedics*, *6*(1), 48. https://doi.org/10.1186/s40634-019-0217-1

Irrgang, J. J., Anderson, A. F., Boland, A. L., Harner, C. D., Kurosaka, M., Neyret, P.,…Shelborne, K. D. (2001). Development and validation of the International Knee Documentation Committee Subjective Knee Form [Article]. *American Journal of Sports Medicine*, *29*(5), 600-613. https://doi.org/10.1177/03635465010290051301

Jansson, H. L., Oji, N. M., Bradley, K. E., Ma, C. B., Zhang, A. L., & Feeley, B. T. (2021). High Rate of Overlapping Question Content Among Commonly Used Patient-Reported Outcome Measurements for Anterior Cruciate Ligament Injury. *Arthrosc Sports Med Rehabil*, *3*(6), e1775-e1801. https://doi.org/10.1016/j.asmr.2021.08.006

Kinikli, G. I., Celik, D., Yuksel, I., & Atay, O. A. (2015). Turkish version of the Anterior Cruciate Ligament Quality of Life questionnaire. *Knee Surg Sports Traumatol Arthrosc*, *23*(8), 2367-2375. https://doi.org/10.1007/s00167-014-3404-8

Kocher, M. S., Smith, J. T., Iversen, M. D., Brustowicz, K., Ogunwole, O., Andersen, J.,…Zurakowski, D. (2011). Reliability, validity, and responsiveness of a modified International Knee Documentation Committee Subjective Knee Form (Pedi-IKDC) in children with knee disorders. *Am J Sports Med*, *39*(5), 933-939. https://doi.org/10.1177/0363546510383002

Kujala, U. M., Jaakkola, L. H., Koskinen, S. K., Taimela, S., Hurme, M., & Nelimarkka, O. (1993). Scoring of patellofemoral disorders. *Arthroscopy*, *9*(2), 159-163. https://doi.org/10.1016/s0749-8063(05)80366-4

Kvist, J., Österberg, A., Gauffin, H., Tagesson, S., Webster, K., & Ardern, C. (2013). Translation and measurement properties of the Swedish version of ACL-Return to Sports after Injury questionnaire. *Scand J Med Sci Sports*, *23*(5), 568-575. https://doi.org/10.1111/j.1600-0838.2011.01438.x

Lafave, M. R., Hiemstra, L., Kerslake, S., Heard, M., & Buchko, G. (2017). Validity, Reliability, and Responsiveness of the Anterior Cruciate Ligament Quality of Life Measure: A Continuation of Its Overall Validation. *Clin J Sport Med*, *27*(1), 57-63. https://doi.org/10.1097/jsm.0000000000000292

Larivière, V., Kiermer, V., MacCallum, C. J., McNutt, M., Patterson, M., Pulverer, B.,…Curry, S. (2016). A simple proposal for the publication of journal citation distributions. *bioRxiv*, 062109. https://doi.org/10.1101/062109

Larsen, P., Rathleff, M. S., Roos, E. M., & Elsoe, R. (2023). Knee injury and osteoarthritis outcome score (KOOS) - National record-based reference values. *Knee*, *43*, 144-152. https://doi.org/10.1016/j.knee.2023.06.004

Letchford, R., Sparkes, V., & van Deursen, R. W. (2015). Assessing participation in the ACL injured population: Selecting a patient reported outcome measure on the basis of measurement properties. *Knee*, *22*(3), 262-269. https://doi.org/10.1016/j.knee.2015.01.010

Lysholm, J., & Gillquist, J. (1982). Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. *Am J Sports Med*, *10*(3), 150-154. https://doi.org/10.1177/036354658201000306

Macchiarola, L., Grassi, A., Di Paolo, S., Pizza, N., Trisolino, G., Stallone, S.,…Zaffagnini, S. (2020). The Italian cross-cultural adaptations of the paediatric International Knee Documentation Committee Score and the Hospital for Special Surgery Paediatric Functional Activity Brief Scale are reliable instruments in paediatric population. *Knee Surg Sports Traumatol Arthrosc*, *28*(8), 2657-2662. https://doi.org/10.1007/s00167-020-05903-y

Martinez-Calderon, J., Infante-Cano, M., Matias-Soto, J., Perez-Cabezas, V., Galan-Mercant, A., & Garcia-Muñoz, C. (2025). The Incidence of Sport-Related

Anterior Cruciate Ligament Injuries: An Overview of Systematic Reviews Including 51 Meta-Analyses. *J Funct Morphol Kinesiol*, *10*(2). https://doi.org/10.3390/jfmk10020174

Martinkėnienė, V. B., Austys, D., Šaikus, A., Brazaitis, A., Bernotavičius, G., Makulavičius, A., & Verkauskas, G. (2023). The Significance of Selecting an Appropriate Patient-Reported Outcome Measure (PROM): A Cross-Cultural Adaptation of the Specific Paediatric International Documentation Committee Subjective (Pedi-IKDC) Knee Form. *Children (Basel)*, *10*(12). https://doi.org/10.3390/children10121930

Meredith, S. J., Rauer, T., Chmielewski, T. L., Fink, C., Diermeier, T., Rothrauff, B. B.,…Lesniak, B. P. (2020). Return to sport after anterior cruciate ligament injury: Panther Symposium ACL Injury Return to Sport Consensus Group. *Knee Surg Sports Traumatol Arthrosc*, *28*(8), 2403-2414. https://doi.org/10.1007/s00167-020-06009-1

Merton, R. K. (1968). The Matthew Effect in Science. *Science*, *159*(3810), 56-63. https://doi.org/10.1126/science.159.3810.56

Mitchell, J., Magnussen, R. A., Collins, C. L., Currie, D. W., Best, T. M., Comstock, R. D., & Flanigan, D. C. (2015). Epidemiology of Patellofemoral Instability Injuries Among High School Athletes in the United States. *Am J Sports Med*, *43*(7), 1676-1682. https://doi.org/10.1177/0363546515577786

Mohtadi, N. (1998). Development and validation of the quality of life outcome measure (questionnaire) for chronic anterior cruciate ligament deficiency [Article]. *American Journal of Sports Medicine*, *26*(3), 350-357. https://doi.org/10.1177/03635465980260030201

Mokkink, L. B., Elsman, E. B. M., & Terwee, C. B. (2024). COSMIN guideline for systematic reviews of patient-reported outcome measures version 2.0. *Quality of Life Research*, *33*(11), 2929-2939. https://doi.org/10.1007/s11136-024-03761-6

Montalvo, A. M., Schneider, D. K., Webster, K. E., Yut, L., Galloway, M. T., Heidt, R. S., Jr.,…Myer, G. D. (2019). Anterior Cruciate Ligament Injury Risk in Sport: A Systematic Review and Meta-Analysis of Injury Incidence by Sex and Sport Classification. *J Athl Train*, *54*(5), 472-482. https://doi.org/10.4085/1062-6050-407-16

Negahban, H., Mostafaee, N., Sohani, S. M., Mazaheri, M., Goharpey, S., Salavati, M.,…Montazeri, A. (2011). Reliability and validity of the Tegner and Marx activity rating scales in Iranian patients with anterior cruciate ligament injury. *Disabil Rehabil*, *33*(23-24), 2305-2310. https://doi.org/10.3109/09638288.2011.570409

Negahban, H., Pouretezad, M., Yazdi, M. J., Sohani, S. M., Mazaheri, M., Salavati, M.,…Salehi, R. (2012). Persian translation and validation of the Kujala Patellofemoral Scale in patients with patellofemoral pain syndrome. *Disabil Rehabil*, *34*(26), 2259-2263. https://doi.org/10.3109/09638288.2012.683480

Noyes, F. R., Barber, S. D., & Mooar, L. A. (1989). A rationale for assessing sports activity levels and limitations in knee disorders. *Clin Orthop Relat Res*(246), 238-249.

Nunnally, J. C. (1978). An Overview of Psychological Measurement. In B. B. Wolman (Ed.), *Clinical Diagnosis of Mental Disorders: A Handbook* (pp. 97-146). Springer US. https://doi.org/10.1007/978-1-4684-2490-4_4

Padua, R., Bondi, R., Ceccarelli, E., Bondi, L., Romanini, E., Zanoli, G., & Campi, S. (2004). Italian version of the International Knee Documentation Committee Subjective Knee Form: cross-cultural adaptation and validation. *Arthroscopy*, *20*(8), 819-823. https://doi.org/10.1016/j.arthro.2004.06.011

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D.,…Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Bmj*, *372*, n71. https://doi.org/10.1136/bmj.n71

Papadopoulos, C., Constantinou, A., Cheimonidou, A. Z., & Stasinopoulos, D. (2017). Greek cultural adaption and validation of the Kujala anterior knee pain scale in patients with patellofemoral pain syndrome. *Disabil Rehabil*, *39*(7), 704-708. https://doi.org/10.3109/09638288.2016.1161834

Paradowski, P. T., Witoński, D., Kęska, R., & Roos, E. M. (2013). Cross-cultural translation and measurement properties of the Polish version of the Knee injury and Osteoarthritis Outcome Score (KOOS) following anterior cruciate ligament reconstruction. *Health Qual Life Outcomes*, *11*, 107. https://doi.org/10.1186/1477-7525-11-107

Piussi, R., Lindskog, J., Hamrin Senorski, R., Thomeé, R., Samuelsson, K., & Hamrin Senorski, E. (2025). Thresholds for minimum clinically important difference, minimal important change and patient acceptable symptom state for the ACL-RSI and the K-SES in patients after anterior cruciate ligament reconstruction. *Knee Surg Sports Traumatol Arthrosc*. https://doi.org/10.1002/ksa.70044

Ra, H. J., Kim, H. S., Choi, J. Y., Ha, J. K., Kim, J. Y., & Kim, J. G. (2014). Comparison of the ceiling effect in the Lysholm score and the IKDC subjective score for assessing functional outcome after ACL reconstruction. *The Knee*, *21*(5), 906-910. https://doi.org/10.1016/j.knee.2014.06.004

Ramalho, A., Duarte-Mendes, P., Paulo, R., & Petrica, J. (2025). The Tempo Giusto: A Call for Recalibrating Time in Human Movement Research. *Quest*, 1-32. https://doi.org/10.1080/00336297.2025.2586504

Raney, E., Onyeukwu, C., Dvorsky, J., Cunningham, M., & Hughes, J. D. (2025). Managing patellofemoral instability in in-season athletes: a review. *Ann Jt*, *10*, 37. https://doi.org/10.21037/aoj-24-54

Roos, E. M., Roos, H. P., Lohmander, L. S., Ekdahl, C., & Beynnon, B. D. (1998). Knee Injury and Osteoarthritis Outcome Score (KOOS) - Development of a self-administered outcome measure [Article]. *Journal of Orthopaedic and Sports Physical Therapy*, *28*(2), 88-96. https://doi.org/10.2519/jospt.1998.28.2.88

Salavati, M., Akhbari, B., Mohammadi, F., Mazaheri, M., & Khorrami, M. (2011). Knee injury and Osteoarthritis Outcome Score (KOOS); reliability and validity in competitive athletes after anterior cruciate ligament reconstruction. *Osteoarthritis Cartilage*, *19*(4), 406-410. https://doi.org/10.1016/j.joca.2011.01.010

Schmitt, L. C., & Rudolph, K. S. (2008). Muscle stabilization strategies in people with medial knee osteoarthritis: The effect of instability. *Journal of Orthopaedic Research*, *26*(9), 1180-1185. https://doi.org/https://doi.org/10.1002/jor.20619

Schrijvers, J. C., van den Noort, J. C., van der Esch, M., Dekker, J., & Harlaar, J. (2019). Objective parameters to measure (in)stability of the knee joint during gait: A review of literature. *Gait & Posture*, *70*, 235-253. https://doi.org/https://doi.org/10.1016/j.gaitpost.2019.03.016

Sharma, L., Chmiel, J. S., Almagor, O., Moisio, K., Chang, A. H., Belisle, L.,…Hayes, K. W. (2015). Knee instability and basic and advanced function decline in knee osteoarthritis. *Arthritis care & research*, *67*(8), 1095-1102.

Silva, L. O., Mendes, L. M. R., Lima, P. O. P., & Almeida, G. P. L. (2018). Translation, cross-adaptation and measurement properties of the Brazilian version of the ACL-RSI Scale and ACL-QoL Questionnaire in patients with anterior cruciate ligament reconstruction. *Braz J Phys Ther*, *22*(2), 127-134. https://doi.org/10.1016/j.bjpt.2017.09.006

Slagers, A. J., Reininga, I. H., & van den Akker-Scheek, I. (2017). The Dutch language anterior cruciate ligament return to sport after injury scale (ACL-RSI) - validity and reliability. *J Sports Sci*, *35*(4), 393-401. https://doi.org/10.1080/02640414.2016.1167230

Sundemo, D., Alentorn-Geli, E., Hoshino, Y., Musahl, V., Karlsson, J., & Samuelsson, K. (2016). Objective measures on knee instability: dynamic tests: a review of devices for assessment of dynamic knee laxity through utilization of the pivot shift test. *Curr Rev Musculoskelet Med*, *9*(2), 148-159. https://doi.org/10.1007/s12178-016-9338-7

Tegner, Y., & Lysholm, J. (1985). Rating systems in the evaluation of knee ligament injuries [Article]. *Clinical Orthopaedics and Related Research*, *198*, 43-49. https://doi.org/10.1097/00003086-198509000-00007

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J.,…Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res*, *27*(5), 1159-1170. https://doi.org/10.1007/s11136-018-1829-0

Tigerstrand Grevnerts, H., Grävare Silbernagel, K., Sonesson, S., Ardern, C., Österberg, A., Gauffin, H., & Kvist, J. (2017). Translation and testing of measurement properties of the Swedish version of the IKDC subjective knee form. *Scand J Med Sci Sports*, *27*(5), 554-562. https://doi.org/10.1111/sms.12861

van der Esch, M., van der Leeden, M., Roorda, L. D., Lems, W. F., & Dekker, J. (2016). Predictors of self-reported knee instability among patients with knee osteoarthritis: results of the Amsterdam osteoarthritis cohort. *Clinical Rheumatology*, *35*(12), 3007-3013. https://doi.org/10.1007/s10067-016-3411-x

van der Velden, C. A., van der Steen, M. C., Leenders, J., van Douveren, F., Janssen, R. P. A., & Reijman, M. (2019). Pedi-IKDC or KOOS-child: which questionnaire should be used in children with knee disorders? *BMC Musculoskelet Disord*, *20*(1), 240. https://doi.org/10.1186/s12891-019-2600-6

van Meer, B. L., Meuffels, D. E., Vissers, M. M., Bierma-Zeinstra, S. M., Verhaar, J. A., Terwee, C. B., & Reijman, M. (2013). Knee injury and Osteoarthritis Outcome Score or International Knee Documentation Committee Subjective Knee Form: which questionnaire is most useful to monitor patients with an anterior cruciate ligament rupture in the short term? *Arthroscopy*, *29*(4), 701-715. https://doi.org/10.1016/j.arthro.2012.12.015

Vincent, W. J., & Weir, J. P. (2012). *Statistics in Kinesiology*. Human Kinetics.

Wang, J. (2014). Unpacking the Matthew effect in citations. *Journal of Informetrics*, *8*(2), 329-339. https://doi.org/https://doi.org/10.1016/j.joi.2014.01.006

Wang, W., Liu, L., Chang, X., Jia, Z. Y., Zhao, J. Z., & Xu, W. D. (2016a). Cross-cultural translation of the Lysholm knee score in Chinese and its validation in patients with anterior cruciate ligament injury. *BMC Musculoskeletal Disorders*, *17*(1), 436. https://doi.org/10.1186/s12891-016-1283-5

Webster, K. E., & Feller, J. A. (2018). Development and Validation of a Short Version of the Anterior Cruciate Ligament Return to Sport After Injury (ACL-RSI) Scale. *Orthopaedic Journal of Sports Medicine*, *6*(4), 2325967118763763. https://doi.org/10.1177/2325967118763763

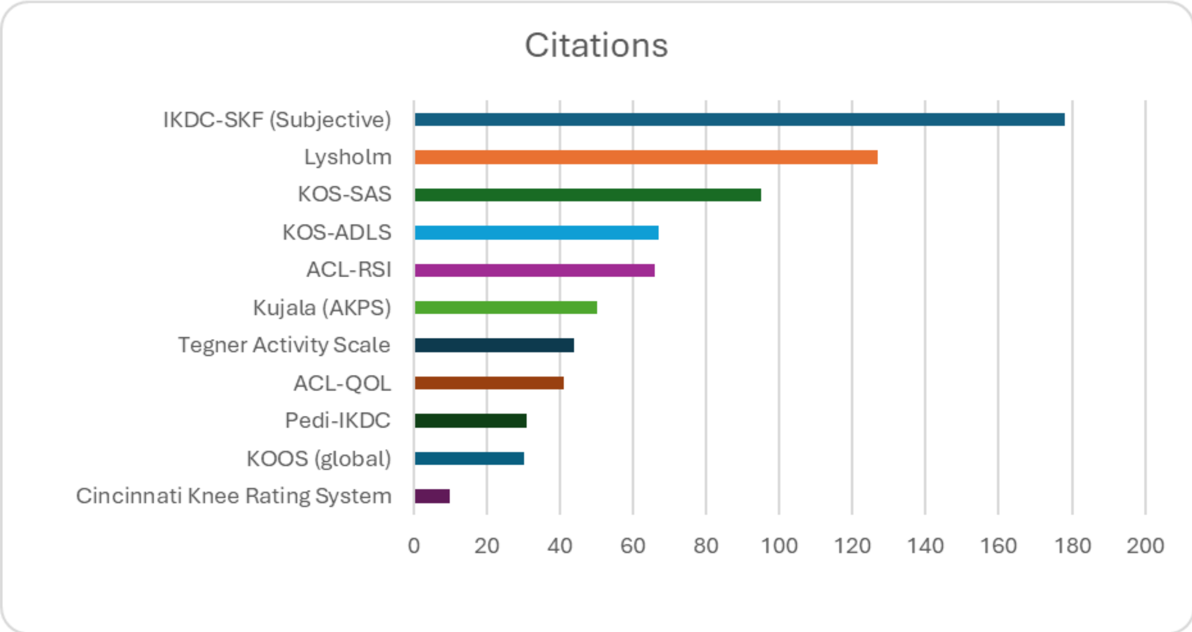
Webster, K. E., Feller, J. A., & Lambros, C. (2008). Development and preliminary validation of a scale to measure the psychological impact of returning to sport following anterior cruciate ligament reconstruction surgery. *Phys Ther Sport*, *9*(1), 9-15. https://doi.org/10.1016/j.ptsp.2007.09.003

Wilsdon, J. (2015). *The Metric Tide: Independent Review of the Role of Metrics in Research Assessment and Management*. SAGE Publications Ltd. https://doi.org/10.4135/9781473978782

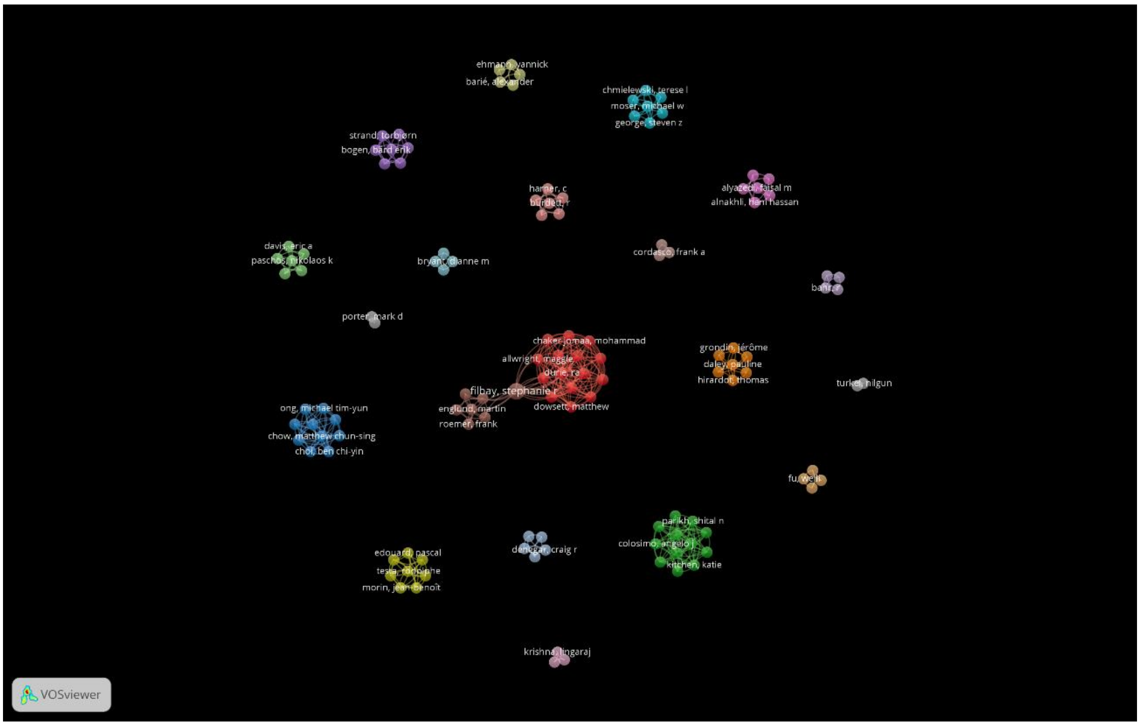
Zlotnicki, J. P., Naendrup, J. H., Ferrer, G. A., & Debski, R. E. (2016). Basic biomechanic principles of knee instability. *Curr Rev Musculoskelet Med*, *9*(2), 114-122. https://doi.org/10.1007/s12178-016-9329-8

SUPPLEMENTARY FIGURES AND TABLES

Supplementary Figure 1. Instruments citations accounts



Supplementary Figure 2. Authors and Co-authors Network



Supplementary Figure 3. Co-authorship analysis based on all author affiliations by countries

Supplementary Figure 4. Organisation-level co-authorship mapping based on author affiliations

Supplementary Table 1. Summary of the Structures That Stabilize the Knee Joint (Abulhasan & Grey, 2017)

Structure	**Type of Stability**	**Innervation**	**Origin**	**Insertion**	**Action**
Postero-Anterior (PA) Stability					
Anterior Cruciate Ligament (ACL)	Primary	Tibial nerve	Anterolateral aspect of the medial femoral condyle	Posterior aspect of the tibial plateau	Restrains postero-anterior tibial translation relative to the femur, tibial external rotation, and knee extension
Rectus Femoris	Secondary	Femoral nerve	Anterior inferior iliac spine and acetabulum	Patellar tendon	Hip flexion and knee extension
Vastus Lateralis	Secondary	Femoral nerve	Greater trochanter and intertrochanteric line of the femur	Patella and tibial tuberosity	Knee extension

Vastus Medialis	Secondary	Femoral nerve	Medial aspect of the femur	Quadriceps tendon	Knee extension
Vastus Intermedius	Secondary	Femoral nerve	Anterolateral aspect of the femur	Quadriceps tendon	Knee extension
Antero-Posterior (AP) Stability					
Posterior Cruciate Ligament (PCL)	Primary	Tibial nerve	Posteromedial aspect of the lateral femoral condyle	Anterior aspect of the tibial plateau	Restrains antero-posterior tibial translation and anterolateral tibial rotation relative to the femur
Biceps Femoris – Long Head	Secondary	Tibial nerve	Ischial tuberosity	Head of the fibula	Knee flexion, lateral rotation of the leg, and hip extension
Biceps Femoris – Short Head	Secondary	Common fibular nerve	Linea aspera of the femur	Head of the fibula	Knee flexion and lateral

					rotation of the leg
Semimembranosus	Secondary	Sciatic nerve	Ischial tuberosity	Medial condyle of the tibia	Knee flexion and hip extension
Semitendinosus	Secondary	Sciatic nerve	Ischial tuberosity	Pes anserinus (tibia)	Knee flexion, hip extension, and ankle plantarflexion
Plantaris	Secondary	Tibial nerve	Lateral supracondylar ridge of the femur	Calcaneal tendon	Knee flexion and ankle plantarflexion
Gastrocnemius	Secondary	Tibial nerve	Medial and lateral femoral condyles	Calcaneus	Knee flexion and ankle plantarflexion
Sartorius	Secondary	Femoral nerve	Anterior superior iliac spine	Pes anserinus (tibia)	Hip flexion, lateral rotation of the thigh, and knee flexion

Varus-Valgus Stability					
Medial Collateral Ligament (MCL)	Primary	Femoral nerve	Medial femoral epicondyle	Periosteum of the proximal tibia	Restrains valgus stress and anteromedial tibial rotation
Lateral Collateral Ligament (LCL)	Primary	Common fibular nerve	Lateral femoral epicondyle	Fibular head	Restrains varus stress, posteromedial tibial rotation, and medial tibial rotation (with the femur fixed)
Popliteus	Secondary	Tibial nerve	Lateral femoral epicondyle	Posterior surface of the tibia	Lateral femoral rotation (with the tibia fixed)
Tensor Fasciae Latae	Secondary	Superior gluteal nerve	Iliac crest	Iliotibial tract	Hip flexion and medio-lateral rotational control of the knee

Supplementary Table 2. Database-specific search strategy

Database	**Search strategy**
PubMed	(("knee"[Title/Abstract] OR "knee joint"[Title/Abstract] OR "knee injury"[Title/Abstract] OR "functional instability"[Title/Abstract]) AND (sport*[Title/Abstract] OR exercise*[Title/Abstract] OR training[Title/Abstract]) AND ("psychometric properties"[Title/Abstract] OR reliability[Title/Abstract] OR validity[Title/Abstract] OR "validation studies"[Title/Abstract] OR "internal consistency"[Title/Abstract]) AND (survey[Title/Abstract] OR questionnaire*[Title/Abstract] OR instrument*[Title/Abstract] OR measur*[Title/Abstract] OR scale[Title/Abstract] OR "subjective evaluation"[Title/Abstract] OR assess*[Title/Abstract] OR tool[Title/Abstract] OR self-report*[Title/Abstract]))
Scopus	TITLE-ABS-KEY(("knee" OR "knee joint" OR "knee injury" OR "functional instability") AND (sport* OR exercise* OR training) AND ("psychometric properties" OR reliability OR validity OR "validation studies" OR "internal consistency") AND (survey OR questionnaire* OR instrument* OR measur* OR scale OR "subjective evaluation" OR assess* OR tool OR self-report*))
Web of Science	Topic=(("knee" OR "knee joint" OR "knee injury" OR "functional instability") AND (sport* OR exercise* OR training) AND ("psychometric properties" OR reliability OR validity OR "validation studies" OR "internal consistency") AND (survey OR questionnaire* OR instrument* OR measur* OR scale OR "subjective evaluation" OR assess* OR tool OR self-report*))
SportDiscus	(("knee" OR "knee joint" OR "knee injury" OR "functional instability") AND (sport* OR exercise* OR training) AND ("psychometric properties" OR reliability OR validity OR "validation studies" OR "internal consistency") AND (survey OR questionnaire* OR instrument* OR measur* OR scale OR "subjective evaluation" OR assess* OR tool OR self-report*))

Supplementary Table 3. Journal-level bibliometric and impact indicators of the included documents

Journal	Documents	%	SCImago category	SJR	Total citations	Quartile	IF 2024	Docs × IF	Weight in sample	Contribution to weighted IF
American Journal of Sports Medicine	3	15%	Physical Therapy, Sports Therapy and Rehabilitation, Sports Science, Medicine (miscellaneous), Orthopedics and Sports Medicine	2.26	6904	Q1	4.5	13.5	15%	0.67
Bmj Open	1	5%	Medicine (miscellaneous)	1.016	36796	Q1	2.3	2.3	5%	0.11
British Journal of Sports Medicine	1	5%	Physical Therapy, Sports Therapy and Rehabilitation, Sports Science, Medicine (miscellaneous), Orthopedics and Sports Medicine	4.724	7484	Q1	16.2	16.2	5%	0.81
Clinical Orthopaedics and Related Research	1	5%	Sports Science, Medicine (miscellaneous), Orthopedics and Sports Medicine, Surgery	1.339	2478	Q1	4.4	4.4	5%	0.22
Journal of Clinical Medicine	1	5%	Medicine (miscellaneous)	0.919	68946	Q1	2.9	2.9	5%	0.14
Journal of Knee Surgery	1	5%	Orthopedics and Sports Medicine, Surgery	0.852	1312	Q1	1.6	1.6	5%	0.08
Journal of Orthopaedic & Sports Physical Therapy	1	5%	Physical Therapy, Sports Therapy and Rehabilitation, Sports Science, Medicine (miscellaneous)	1.606	1297	Q1	5.8	5.8	5%	0.29
Journal of Orthopaedic Surgery and Research	1	5%	Orthopedics and Sports Medicine, Surgery	0.877	7159	Q1	2.8	2.8	5%	0.14
Journal of Pediatric Orthopaedics	1	5%	Medicine (miscellaneous), Orthopedics and Sports Medicine, Pediatrics, Perinatology and Child Health	0.655	1418	Q2	1.5	1.5	5%	0.08
Knee Surgery Sports Traumatology Arthroscopy	4	20%	Sports Science, Orthopedics and Sports Medicine, Surgery	1.95	7594	Q1	5	20	20%	1.00
Orthopaedic Journal of Sports Medicine	3	15%	Orthopedics and Sports Medicine	1.258	4775	Q1	2.5	7.5	15%	0.38
Scandinavian Journal of Medicine & Science in Sports	1	5%	Physical Therapy, Sports Therapy and Rehabilitation, Sports Science, Orthopedics and Sports Medicine	1.494	2730	Q1	3.8	3.8	5%	0.19
Trials	1	5%	Medicine (miscellaneous), Pharmacology (medical)	0.893	5818	Q2	2	2	5%	0.10

Supplementary Table 4. Detailed COSMIN Risk-of-Bias Assessment by Study

PROM	Study	Internal consistency	Reliability	Measurement error	Content validity	Structural validity	Hypotheses testing	Cross-cultural adaptation	Cross-cultural validity	Responsiveness
ACL-QOL	Mohtadi (1998)		F?		G+		F+			F+
ACL-QOL	Bryant et al. (2006)		P+	F?						
ACL-QOL	Kinikli et al. (2014)	F?	F+	F?	F?		F+	G+		F?
ACL-QOL	Lafave et al. (2016)	F?	F?	F?						F?
ACL-RSI	Webster et al. (2008)	F?			F?	F?	F+			
ACL-RSI	Slagers et al. (2016)	F?	F+	F?	F?	F?	F+	G+		
ACL-RSI	Kvist et al. (2012)	F?	F+	F?	F?	F?	F+	G+		
AKPS / Kujala	da Cunha et al., 2013	F?	F+	F?	F?		F+	G+		F+
AKPS / Kujala	Buckinx et al., 2017	F?	F+	F?	F?		F+	G+		
AKPS / Kujala	Papadopoulos et al., 2016	F?	F+	F?	F?		F+	F+		
AKPS / Kujala	Negahban et al., 2012	F?	F+	F?	F?		F+	G+		
AKPS / Kujala	da Silva-Júnior et al., 2024					G+				
AKPS / Kujala	Dammerer et al., 2018	F?	F+		F?	P+	F+	G+		F+
Cincinnati Knee Rating System (CKRS)	Barber-Westin et al. (1999)		G+		F+		F+			F+
Cincinnati Knee Rating System (CKRS)	Hopper et al. (2002)		P+							
IKDC (activity item)	Letchford et al. (2015)		F+	F+	F?		F+			F+
IKDC-SKF	Çelik et al. (2014)	F?	P+	F?	F?		F+	G+		F?
IKDC-SKF	Padua et al. (2004)	F?	P+		F?		F?	F+		
IKDC-SKF	Ra et al., 2014				F?		F?			
IKDC-SKF	Irrgang et al., 2001	F?	F+	F?	F?	F?	F+			
IKDC-SKF	van Meer et al. (2013)		F+	F?	F+		F+			F+
IKDC-SKF	Ebrahimzadeh et al. (2014)	F?	P+		F?		F?	G+		
KOOS	van Meer et al. (2013)		G+	F?	F?		F?			F?
KOOS	Roos et al. (1998)		P+		F?		F?			F?
KOOS	Comins et al. (2007)	F?				F?				
KOOS	Salavati et al. (2011)	F?	P+	F?			F+			
KOOS	Paradowski et al. (2013)	F?	F+	F?	F?		F+	G+		F+
Lysholm knee scoring scale	Briggs et al. (2009)	F?	G+	F?	F?		G+			G+
Lysholm knee scoring scale	Bengtsson et al. (1996)		P+				F?			
Lysholm knee scoring scale	Ra et al., 2014				F?		F?			
Pedi-IKDC	van der Velden et al., 2019		F+	F?	G+		F+	G+		F+
Pedi-IKDC	Macchiarola et al., 2020	F?	F+	F?	F?		F?	G+		F?
Pedi-IKDC	Martinkėnienė et al., 2023	F?	F+	F?	F?		F?	G+		F?
Tegner Activity Scale	Briggs et al. (2009)		F+	F?	F?		F+			G+
Tegner Activity Scale	Negahban et al. (2011)		F+	F?	F?		F+	G+		
Tegner Activity Scale	Huang et al. (2016)		F+	F?	F?		F+	G+		
Tegner Activity Scale	Letchford et al. (2015)		F+	F?	F?		F+			F+

Note 3. Abbreviations / codes: G+ = sufficient evidence with good methodological quality; F+ = sufficient evidence with fair methodological quality; P+ = sufficient evidence with poor methodological quality; F? = evidence present but insufficient for a definitive sufficiency judgement; ? = unclear or indeterminate; blank cell = no extractable evidence retained in the audited matrix.

Supplementary Table 5. Study-Level Extracted Psychometric Metrics and PROM-Specific Notes

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
da Silva-Júnior et al. (2024)	AKPS / Kujala	NR	NR	NR	NR	NR	NR	NR	13-item: 0.059 11-item: 0.067	r=0.966; p<0.001 (13-item vs 11-item)	NR	NR	NR	NR	NR	CFA/RDWLS. 13-item model fit adequate but items 11–12 had low loadings (~0.22). 11-item model retained adequate fit and lower AIC/SABIC. No reliability, measurement error or responsiveness metrics reported.
Dammerer et al. (2018)	AKPS / Kujala	0.87	0.93	NR	NR	NR	6.1; 1.2	46.7%; 9.2%	NR	KOOS Symptoms r=0.65; Pain r=0.78; ADL r=0.74; Sport/Rec r=0.84; QOL r=0.70; Lysholm r=0.88; VAS pain r=-0.71; SF-12 PCS r=0.79; SF-12 MCS r=0.14	d=0.57 (pre-op to 6 mo); d=0.96 (pre-op to 12 mo)	NR	10 points (literature-based; not estimated)	NR	NR	German translation/validation of the Kujala score in patients with patellofemoral instability undergoing MPFL reconstruction. The study reported internal consistency (alpha=0.87), test-retest reliability after 1 week (ICC=0.93), exploratory PCA with two factors (eigenvalues 6.1 and 1.2; 46.7% and 9.2% variance), strong convergent validity with KOOS/Lysholm/VAS /SF-12 PCS, weak correlation with SF-12 MCS supporting discriminant validity, and responsiveness by

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																Cohen's d at 6 and 12 months. No SEM, SDC/MDC, SRM, MCID or AUC were reported. MIC=10 points was assumed from prior literature for sample-size planning, not estimated in this study.
Martinkė nienė et al. (2023)	Pedi-IKDC	Total 0.91; Symptoms 0.75; Sport/function 0.92	Overall 0.98 (95% CI 0.97 to 0.99); item ICCs 0.87–0.98	2.32 (Table 3; abstract reports 2.97)	8.23	NR	NR	NR	NR	Pedi-IKDC total: Lysholm r=0.78; PedsQL total r=0.27; PedsQL physical functioning r=0.65. Symptoms: Lysholm r=0.72; PedsQL total r=0.26; PedsQL physical functioning r=0.57. Sport/function: Lysholm r=0.73; PedsQL total r=0.26; PedsQL physical functioning r=0.67.	1.98	1.72	NR	NR	NR	Lithuanian cross-cultural adaptation and validation study of the Pedi-IKDC in 60 paediatric patients with various knee disorders (57 retest; 42 responsiveness). Cronbach's alpha was acceptable for the total score and components. Test–retest reliability was excellent (overall ICC=0.98), with item ICCs ranging from 0.87 to 0.98. Table 3 reports SEM=2.32 and SDC=8.23; the abstract reports SEM=2.97, which is mathematically consistent with SDC=8.23, so the discrepancy is documented. Criterion validity was assessed against Lysholm and PedsQL; correlations were moderate with Lysholm and PedsQL physical functioning,

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																but weak with total PedsQL. Floor and ceiling effects were 3.3% and 1.6%. Responsiveness was large (ES=1.98; SRM=1.72). No structural validity analysis, MIC/MCID or AUC were reported.
Macchiarola et al. (2020)	Pedi-IKDC	0.92	0.96	4.4	12.3	NR	NR	NR	NR	Pedi-FABS overall r=0.60; item-level correlations: Pedi-FABS items 1–6 moderate (0.3<r<0.7), items 7–8 low (r<0.3).	0.79	0.86	NR	NR	NR	Italian cross-cultural adaptation and validation study of the Pedi-IKDC and Pedi-FABS in 89 paediatric patients with knee pathologies (81 retest; 49 responsiveness). Only Pedi-IKDC is retained in this S5 because Pedi-FABS is outside the predefined 9 PROMs. For Pedi-IKDC, reliability and measurement error were reported as ICC=0.96, SEM=4.4 and SDC=12.3. Internal consistency was high (Cronbach's alpha=0.92). Criterion/convergent validity was assessed against Pedi-FABS, with moderate-to-low correlations (overall r=0.60). Floor and ceiling effects were 0% and 22%, respectively. Responsiveness was distribution-based,

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenv alue	% varian ce	RMS EA	Correlation r	Effect size	SRM	MIC	MC ID	AUC	Key extracted metrics
																with ES=0.79 and SRM=0.86. No structural validity analysis, MDC, MIC/MCID or AUC were reported.
van der Velden et al. (2019)	Pedi-IKDC	NR	0.90	8.6	23.8	NR	NR	NR	NR	KOOS-Child Symptoms r=0.69; Pain r=0.84; ADL r=0.83; Sport/play r=0.80; QOL r=0.76; Lysholm r=0.71; NRS Pain r=-0.70; EQ-5D r=0.53.	No change ES=0.24; improvement ES=1.50; worsening ES=-0.66	No change SRM=0.20; improvement SRM=1.25; worsening SRM=-0.41	NR	NR	NR	Dutch translation and psychometric comparison of Pedi-IKDC and KOOS-Child in children with knee disorders. Only Pedi-IKDC is retained in this S5 because KOOS-Child is outside the predefined 9 PROMs. Pedi-IKDC showed test-retest reliability ICC=0.90, SEM=8.6 and SDC=23.8 in stable patients. Content relevance was 20/21 items (95%) according to both orthopaedic surgeons and patients. Construct validity confirmed 6/8 predefined hypotheses (75%), with correlations against KOOS-Child subscales, Lysholm, NRS Pain and EQ-5D. Interpretability showed floor=0% and ceiling=1%. Responsiveness confirmed 3/4 hypotheses (75%), with ES/SRM values reported separately for

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																no change, improvement and worsening groups. Cronbach alpha, structural validity indices, MDC, MIC/MCID and AUC were not reported.
Buckinx et al. (2017)	AKPS / Kujala	0.87	0.97 (95% CI 0.96-0.98)	2.58	NR	7.14	NR	NR	NR	SF-36 PF r=0.80; RP r=0.60; BP r=0.74; GH r=0.57; RE r=0.07; VT r=0.32; MH r=0.26; SF r=0.41	NR	NR	NR	NR	NR	French cross-cultural adaptation and validation study of the AKPS/Kujala in 101 French-speaking patients with patellofemoral pain. The F-AKPS showed high internal consistency (Cronbach alpha=0.87), excellent test-retest reliability over 7 days (ICC=0.97, 95% CI 0.96-0.98), SEM=2.58 and MDC=7.14. Construct validity was supported by stronger correlations with SF-36 physical domains (PF, RP, BP, GH) and weaker/no correlations with mental/emotional domains (RE, VT, MH, SF). No floor or ceiling effects were observed. No SDC, eigenvalue, explained variance, RMSEA, responsiveness indices, MIC/MCID or AUC were reported.

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
da Cunha et al. (2013)	AKPS / Kujala	0.75 (alpha if item deleted 0.70-0.77)	0.95 (95% CI 0.91-0.97)	2.98 (2.9%)	NR	NR	NR	NR	NR	Baseline: FIQ r=0.66; PSS r=-0.63; NPRS r=-0.26. Change scores: FIQ r=0.41; PSS r=-0.52; NPRS r=-0.40. External responsiveness vs GPE: r=0.27	1.96 (84% CI 1.07-2.85)	NR	NR	NR	0.70	Brazilian Portuguese cross-cultural adaptation and clinimetric testing study in patients with patellofemoral pain syndrome. Only the AKPS/Kujala line is retained here because FIQ, PSS, NPRS and GPE are outside the nine retained PROMs/scales. For AKPS, the study reported adequate internal consistency (Cronbach alpha=0.75), excellent test-retest reliability over 48-72 h (ICC2,1=0.95, 95% CI 0.91-0.97), and very good agreement (SEM=2.98; 2.9% of total score). Construct validity showed moderate correlations with FIQ (r=0.66) and PSS (r=-0.63), and a weaker correlation with NPRS (r=-0.26). No floor or ceiling effects were observed. Internal responsiveness was high (effect size=1.96, 84% CI 1.07-2.85), and external responsiveness reached the minimum acceptable ROC threshold (AUC=0.70). No

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																SDC, MDC, structural-validity indices, SRM, MIC or MCID were reported.
Papadopoulos et al. (2016)	AKPS / Kujala	0.942	0.921 (95% CI 0.857-0.927)	4.54	6.39	NR	NR	NR	NR	PFPS severity scale: first KAKPS r=0.714; second KAKPS r=0.761; both p<0.001	NR	NR	NR	NR	NR	Greek cross-cultural adaptation and validation study of the AKPS/Kujala anterior knee pain scale (KAKPS) in 130 adults aged 18-45 years with anterior knee pain/patellofemoral pain syndrome for at least 4 weeks. The KAKPS showed high internal consistency (Cronbach alpha=0.942), excellent test-retest reliability in 112 stable patients over 2-3 days (ICC=0.921, 95% CI 0.857-0.927), SEM=4.54 and SDC=6.39. Concurrent validity was supported by significant correlations with the PFPS severity scale at the first and second administrations (r=0.714 and r=0.761, p<0.001). No floor or ceiling effects were observed. No structural-validity indices, responsiveness metrics, MIC/MCID

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																or AUC were reported.
Negahban et al. (2012)	AKPS / Kujala	0.81	0.96 (95% CI 0.93-0.98)	NR	6.44	NR	NR	NR	NR	SF-36: PF 0.51; RP 0.44; BP 0.47; GH 0.34; VT 0.33; SF 0.37; RE 0.25; MH 0.35	NR	NR	NR	NR	NR	Persian translation and validation of the AKPS/Kujala Patellofemoral Scale in 100 patients with chronic patellofemoral pain syndrome; retest n=47 after 2-3 days. The Persian KPS showed Cronbach alpha=0.81, ICC2,1=0.96 (95% CI 0.93-0.98), and SDC=6.44. Construct validity was assessed through Spearman correlations with SF-36 subscales: physical-component correlations were generally higher than mental-component correlations (PF 0.51; RP 0.44; BP 0.47; GH 0.34; VT 0.33; SF 0.37; RE 0.25; MH 0.35). No floor or ceiling effects were observed. Responsiveness, MIC/MCID, AUC, factor analysis and RMSEA were not reported.
Letchford et al. (2015)	Tegner Activity Rating Scale	NA	Individual ICC=0.92; group ICC=0.92	0.63	Individual SDC=1.75; group SDC=0.29	NR	NR	NR	NR	Convergent/divergent validity: CSAS r=0.628; Marx r=0.686; IKDC activity item r=-0.719; Lysholm r=0.612; IKDC-	Known-groups ES: PI-PO=0.84; PO-6m=0.52; 6-12m=0.64	NR	ROC MIC cut-off: better +0.5; worse -0.5;	NR	Better: AUC 0.637 minimally, 0.699 much; worse:	Comparative COSMIN-based measurement-property study of four participation PROMs after ACL reconstruction. For the Tegner Activity

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
										SKF r=0.700; Pain VAS r=-0.436; age r=-0.251; BMI r=0.060	. Responsiveness contrasts ES: 1v2=0.80; 1v3=0.86; 2v3=0.18		interpreted clinically as 1-point category change		AUC 0.930 minimally, 0.931 much	Rating Scale, test-retest reliability was ICC=0.92 for both individual and group analyses, SEM=0.63, SDC=1.75 for individual interpretation and 0.29 for group interpretation. Construct validity was examined through correlations with participation PROMs, knee function, pain and demographics. Known-groups validity and responsiveness were supported by longitudinal changes from pre-injury, pre-operative, 6-month and 12-month states. Tegner had no relevant floor/ceiling effects and was the preferred participation PROM. Cronbach alpha and structural-validity indices are not applicable/reported because Tegner is a single-item ordinal activity scale.
Letchford et al. (2015)	IKDC activity item / participation item	NA	Individual ICC=0.86; group ICC=0.86	0.29	Individual SDC=0.80; group SDC=0.14	NR	NR	NR	NR	Convergent/divergent validity: Tegner r=-0.719; CSAS r=-0.528; Marx r=-0.640; Lysholm r=-	Known-groups ES: PI-PO=0.88; PO-6m=0.51; 6-	NR	ROC MIC cut-off: better +0.5; worse	NR	Better: AUC 0.603 minimally, 0.681 much;	Comparative COSMIN-based measurement-property study of four participation PROMs after ACL reconstruction. The

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
										0.567; IKDC-SKF r=-0.710; Pain VAS r=0.400; age r=0.396; BMI r=0.223	12m=0.56. Responsiveness contrasts ES: 1v2=0.83; 1v3=0.83; 2v3=0.03		-0.5; interpreted clinically as 1-point category change		worse: AUC 0.906 minimally, 0.877 much	IKDC result refers to the single IKDC activity/participation item, not to the full IKDC-SKF total score. Test-retest reliability was ICC=0.86 for individual and group analyses, SEM=0.29, SDC=0.80 for individual interpretation and 0.14 for group interpretation. Construct validity was examined through correlations with other participation PROMs and function/pain measures. Known-groups validity and responsiveness were supported, but floor/ceiling effects were present in several time points and sensitivity to minimal improvement was lower than for deterioration. Cronbach alpha and structural-validity indices are not applicable/reported for this single item.
Slagers et al. (2016)	ACL-RSI	0.94	0.93; 95% CI 0.88-0.96	5.5	Individual SDC= 15.3; group SDC= 1.5	NR	7.1	59%	NR	I-PRRS r=0.79; KOOS Pain r=0.36; KOOS Symptoms r=0.30; KOOS ADL r=0.25; KOOS	NR	NR	NR	NR	NR	Dutch translation and validation study of the ACL-RSI in 150 patients 3-16 months after ACL reconstruction, with 107 completing retest

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
										Sport/Rec r=0.48; KOOS QOL r=0.40; IKDC 2000 r=0.51; TSK r=-0.46; MHLC-C r=-0.15						and 104 included in test-retest analysis. PCA supported a one-factor structure (eigenvalue=7.1; explained variance=59%). Nine of ten predefined hypotheses were confirmed (90%). Internal consistency was high (Cronbach alpha=0.94) and test-retest reliability was excellent (ICC=0.93; 95% CI 0.88-0.96). SEM=5.5, SDCind=15.3 and SDCgroup=1.5. Bland-Altman showed a systematic bias of 3.2 points between test and retest, but this was below the SEM. No responsiveness, MIC/MCID or AUC were reported.
Kvist et al. (2012)	ACL-RSI	0.948	0.893	0.7	Individual SDC= 1.9; group SDC= 0.3	NR	7.7	64.1%	NR	Global knee function r=0.713; satisfaction with activity level r=0.486; symptom satisfaction r=-0.561; KOOS Symptoms r=0.479; KOOS Pain r=0.510; KOOS ADL r=0.484; KOOS Sport r=0.594; KOOS QOL	NR	NR	NR	NR	NR	Swedish translation and measurement-property study of the ACL-RSI in 182 patients 2-5 years after ACL reconstruction, with 48 stable patients included in reproducibility analysis. PCA supported a one-factor structure (eigenvalue=7.7; explained variance=64.1%).

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
										r=0.718; ACL-QOL total r=0.815; symptoms r=0.607; work r=0.462; participation r=0.852; lifestyle r=0.753; social/emotional r=0.726; TSK r=-0.689; K-SES r=0.705; I-HLC-C r=0.287; risk behavior r=0.321						Cronbach alpha=0.948. All nine predefined construct-validity hypotheses were confirmed. Test-retest reproducibility was high (ICC=0.893), with SEM=0.7, SDCind=1.9 and SDCgroup=0.3 on the 1-10 ACL-RSI scale. No responsiveness, MIC/MCID or AUC were reported.
Webster et al. (2008)	ACL-RSI	0.96 (results section; abstract reports 0.92)	NR	NR	NR	NR	8.14	67.8%	NR	NR; known-groups validity: returned to sport mean=70 vs not returned mean=46, p<0.0001; adjusted group means: given up sport=39.1, plan to return=54.9, training only=63.0, full competition=76.3; ANCOVA F(3,215)=34.39, p<0.0001	NR	NR	NR	NR	NR	Development and preliminary validation study of the 12-item ACL-RSI in 220 athletes 8-22 months after ACL reconstruction. Items were generated around emotions, confidence in performance and risk appraisal. Internal consistency was high (Cronbach alpha=0.96 in the Results section; abstract reports 0.92). PCA supported a one-factor solution (eigenvalue=8.14; 67.8% variance), so a single total score was used. Known-groups validity was supported: participants who returned to sport

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																scored higher than those who had not returned, and adjusted scores increased from given up sport to full competition. No test-retest reliability, SEM, SDC/MDC, responsiveness, MIC/MCID or AUC were reported.
Hopper et al. (2002)	CKRS	NR	0.97	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	Test-retest reliability study in 19 participants assessed approximately 12 months after ACL reconstruction using semitendinosus graft. CKRS was completed on two occasions separated by 7 days. Mean CKRS scores were 82.1 (SD 9.9; range 66.8-98.9) at occasion 1 and 82.6 (SD 10.4; range 65.8-98.9) at occasion 2, with no significant between-occasion difference (p=0.53). Test-retest reliability was high (ICC=0.97). The study also reported an analogue knee-rating scale (ICC=0.96) and functional hop tests, but these were not retained as separate PROM rows. No Cronbach alpha, SEM, SDC/MDC, structural validity,

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																construct validity, responsiveness, MIC/MCID or AUC were reported for CKRS.
Barber-Westin et al. (1999)	CKRS	NR	Uninjured: 0.71-1.00; patients: 0.68-0.98 (most >=0.75; stair climbing in patients=0.68)	NR	NR	NR	NR	NR	NR	Item-discriminant validity: 59/63 comparisons statistically dissimilar; coefficients mostly -0.28 to +0.28, with few exceptions (e.g., overall score vs AP displacement r=-0.38)	Pain 1.40; swelling 1.18; partial giving way 1.87; full giving way 1.49; symptoms avg 1.74; ADL avg 0.69; sports function avg 1.91; overall score 3.49	Pain 1.18; swelling 1.07; partial giving way 1.70; full giving way 1.44; symptoms avg 1.56; ADL avg 0.72; sports function avg 1.37; overall score 2.48	NR	NR	NR	Reliability/validity/responsiveness study of the Cincinnati Knee Rating System in 350 subjects. Test-retest reliability was assessed in 100 subjects (50 uninjured; 50 patients with chronic knee disorders) over a mean 7-day interval. Validity and responsiveness were assessed in 250 ACL-reconstructed patients followed for at least 2 years. ICCs were high for most CKRS variables; content validity showed no floor/ceiling effects for the overall rating score preoperatively and limited ceiling effect at follow-up (9%). Construct validity was supported by 8/9 clinical hypotheses. Responsiveness was high, with large SRM/effect-size values for most categories; no Cronbach alpha, SEM, SDC/MDC, structural validity,

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																MIC/MCID or AUC were reported.
Huang et al. (2016)	Tegner Activity Score / Tegner-C	NA	Control 0.97 (95% CI 0.93-0.99); ACL-patient group 0.90 (0.84-0.94); pre-ACLR 0.71 (0.41-0.87); 2-3 mo post-ACLR 0.95 (0.87-0.98); 3-12 mo post-ACLR 0.99 (0.97-1.00)	Control 0.16; ACL-patient group 0.50; pre-ACLR 0.77; 2-3 mo post-ACLR 0.32; 3-12 mo post-ACLR 0.16	Control 0.43; ACL-patient group 1.38; pre-ACLR 2.12; 2-3 mo post-ACLR 0.89; 3-12 mo post-ACLR 0.44	NR	NR	NR	NR	IKDC-SKF-C overall: control r=0.56; ACL-patient r=0.84; pre-ACLR r=0.66; 2-3 mo post-ACLR r=0.81; 3-12 mo post-ACLR r=0.66. ACL-patient item-level correlations: Q1 0.74; Q2 0.44; Q3 0.40; Q4 0.56; Q5 0.72; Q6 0.19; Q7 0.80; Q8 0.77; Q9 0.81; Q10 0.06	NR	NR	NR	NR	NR	Simplified-Chinese translation, cross-cultural adaptation and validation of the Tegner Activity Score in 78 participants: healthy controls, pre-ACLR, 2-3 months after ACLR and 3-12 months after ACLR. The retest interval was 5 +/- 2 days. Tegner-C showed high test-retest reliability in controls and post-ACLR groups, but lower reliability in the pre-ACLR group. SEM and SDC were reported by subgroup. No floor or ceiling effects were observed. Construct validity was supported by correlations with IKDC-SKF-C, especially in the ACL-patient group, and by known-groups discrimination between controls, pre-ACLR, early post-ACLR and later post-ACLR groups. Responsiveness, MIC/MCID and AUC were not assessed.
Negahban et al. (2011)	Tegner Activity	NA	0.81; 95% CI	NR	0.75	NR	NR	NR	NR	KOOS Pain 0.34; KOOS Symptoms 0.10;	NR	NR	NR	NR	NR	Persian cross-cultural adaptation and validation study of the

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
	Rating Scale		0.68-0.89							KOOS ADL 0.17; KOOS Sport/Rec 0.19; KOOS QOL -0.07; SF-12 PCSM 0.29; SF-12 MCSM 0.12						Tegner and Marx activity rating scales in 100 patients with ACL injury. Only Tegner was retained for this S5 because Marx is outside the 9 selected scales. Tegner was re-administered to 45 stable patients after 2-6 days. Test-retest reliability was acceptable/high (ICC=0.81; 95% CI 0.68-0.89), with SDC=0.75. Internal consistency was not applicable because Tegner is a single-item activity scale. Construct validity was assessed through correlations with KOOS and SF-12; the highest association was with KOOS Pain (r=0.34), and the association with SF-12 physical component was higher than with the mental component. Ceiling and floor effects were both 0%. No SEM value, structural validity indices, responsiveness, MIC/MCID or AUC were reported.
Briggs et al. (2009)	Lysholm Knee Score	0.72	0.94 (95% CI	3.2	NR	8.9	NR	NR	NR	Criterion validity: IKDC r=0.78; SF-12	6 mo 1.0; 9 mo 1.0; 12 mo	6 mo 0.925; 9 mo 1.1;	NR	NR	NR	Patient-administered validation of the Lysholm score in

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
			0.88-0.96)							physical r=0.43; SF-12 mental r=0.07. Construct validity: ADL difficulty rho=0.47; work difficulty rho=0.43; sports difficulty rho=0.33; all predefined hypotheses significant.	1.1; 24 mo 1.1	12 mo 1.2; 24 mo 0.93				ACL-injured patients. Test-retest reliability was assessed in 50 patients at >=24 months postoperatively with retest within 4 weeks and no health-status change. Internal consistency was acceptable (alpha=0.72). SEM=3.2 and minimum detectable change=8.9 were reported. Criterion validity was supported by correlations with IKDC and SF-12 physical component. Construct validity was supported because all 6 predefined hypotheses were significant. Responsiveness was large at 6, 9, 12 and 24 months. SDC, structural validity metrics, MIC/MCID and AUC were not reported.
Briggs et al. (2009)	Tegner Activity Scale	NA	0.82 (95% CI 0.66-0.89)	0.64	NR	1	NR	NR	NR	Criterion validity: SF-12 physical rho=0.20; SF-12 mental rho=-0.01; IKDC rho=0.22. Construct validity: ADL difficulty rho=0.37; work	6 mo 0.74; 9 mo 1.1; 12 mo 1.0; 24 mo 1.0	6 mo 0.61; 9 mo 0.84; 12 mo 0.96; 24 mo 1.0	NR	NR	NR	Patient-administered validation of the Tegner Activity Scale in ACL-injured patients. Test-retest reliability was acceptable (ICC=0.82; 95% CI 0.66-0.89), SEM=0.64 and minimum detectable change=1. Content

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
										difficulty rho=0.38; sports difficulty rho=0.35; all predefined hypotheses significant.						validity was supported by acceptable floor (8%) and ceiling (3%) effects. Criterion validity showed low but significant associations with SF-12 physical component and IKDC, consistent with Tegner measuring activity level rather than knee symptoms/function. All 6 construct-validity hypotheses were significant. Responsiveness was moderate at 6 months and large from 9 to 24 months. Cronbach alpha is not applicable because Tegner is a single-item activity scale.
Bengtsson et al. (1996)	Lysholm Knee Score	NR	NR	NR	NR	NR	NR	NR	NR	Test-retest Kendall correlations for total Lysholm score: day 1 vs day 3 r=0.75; day 1 vs day 14 r=0.69; day 3 vs day 14 r=0.68; all p<0.0001. Sensitivity/discriminative results: ACL mean 81.1; MT 62.3; PFPS 73.4; LAS 64.6; ACL differed from the other	NR	NR	NR	NR	NR	Reliability/sensitivity study of the 1985 Lysholm Knee Score in 31 patients with ACL rupture, meniscus tear, patellofemoral pain syndrome or lateral ankle sprain. The scale was administered by telephone on days 1, 3 and 14. Reliability was reported as Kendall correlation coefficients, not ICC: total-score correlations were

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenv alue	% varian ce	RMS EA	Correlation r	Effect size	SRM	MIC	MC ID	AUC	Key extracted metrics
										three groups (p<0.01).						0.75, 0.69 and 0.68 across the three testing-occasion comparisons. No Cronbach alpha, SEM, SDC/MDC, structural validity metrics, responsiveness indices, MIC/MCID or AUC were reported. The authors concluded that the Lysholm was less sensitive for ACL patients than for the other diagnostic groups because the ACL group had the highest score.
Lafave et al. (2016)	ACL-QOL	Preop 0.93; 6 mo 0.95; 12 mo 0.96; 24 mo 0.98	0.60	6.16	NR	12.1	NR	NR	NR	Anchor-based responsiveness: ACL-QOL change vs 7-point global rating scale r=0.47; 95% CI 0.07-0.87; p<0.05	Eta squared= 0.61	NR	NR	NR	NR	Prospective validation/continuation study of the ACL-QOL in 579 ACL-deficient patients undergoing ACL reconstruction. ACL-QOL was completed preoperatively and at 6, 12 and 24 months. Internal consistency was high across timepoints (alpha 0.93-0.98). Test-retest reliability in a subset of 31 patients was moderate (ICC[2,k]=0.60), with SEM=6.16 and a 95% CI/minimum detectable change estimate of 12.1 points.

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																Responsiveness was supported by significant serial improvement across preoperative, 6-, 12- and 24-month scores (eta squared=0.61) and by correlation between ACL-QOL change and a 7-point global rating of improvement (r=0.47). No factor analysis/structural validity, SDC, MIC/MCID or AUC were reported.
Kinikli et al. (2014)	ACL-QOL	Overall 0.95; domains: symptoms/physical complaints 0.80; work-related concerns 0.77; recreational activity/sport 0.92; lifestyle 0.82; social/emotional 0.80	Overall 0.98; domains 0.95-0.98	3.1	NR	8.7	NR	NR	NR	KOS-ADLS symptoms 0.03; function 0.18; total 0.14; Lysholm 0.23; SF-36 PF 0.33; RP 0.72; BP 0.34; GH 0.29; VT 0.42; SF 0.58; RE 0.54; MH 0.41; PCS 0.51; MCS 0.62	Overall ACL-QOL: 16 wk 2.1; 2 yr 1.1. Domains at 16 wk: 1.4-1.9; at 2 yr: 0.0-2.0	NR	NR	NR	NR	Turkish cross-cultural adaptation and validation of the ACL-QOL in patients after ACL reconstruction. Internal consistency was high for the overall score (Cronbach alpha=0.95) and acceptable across domains (0.77-0.92). Test-retest reliability was excellent for the overall score (ICC=0.98) and domains (0.95-0.98), with SEM=3.1 and MDC95=8.7. Construct/criterion validity was assessed against KOS-ADLS, Lysholm and SF-36; correlations were weak with KOS-

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																ADLS and Lysholm, but moderate to very good with SF-36 components. Overall floor/ceiling effects were acceptable (floor 0%; ceiling 6.2%), although work-related concerns showed a ceiling effect of 22.9%. Responsiveness was large for the overall ACL-QOL at 16 weeks (ES=2.1) and 2 years (ES=1.1). No structural validity metrics, SRM, MIC/MCID or AUC were reported.
Bryant et al. (2006)	ACL-QOL	NR	0.86; 95% CI 0.75-0.91	6.12; 95% CI 5.20-7.44	NR	NR	NR	NR	NR	Recalled vs actual baseline ACL-QOL: Pearson r=0.88; regression coefficient=0.75 (95% CI 0.65-0.86), p<0.001; mean difference=-3.24 (95% CI -5.43 to -1.04), p=0.01	NR	NR	NR	NR	NR	Methodological recall-validity/agreement study in patients undergoing knee surgery. For the ACL-QOL subgroup, patients undergoing ACL reconstruction completed actual preoperative ACL-QOL and recalled preoperative ACL-QOL approximately 2 weeks postoperatively. Agreement between actual and recalled ACL-QOL was excellent (ICC=0.86; 95% CI 0.75-0.91), with SEM=6.12. Predictive validity

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																was high (Pearson r=0.88; regression coefficient=0.75, 95% CI 0.65-0.86). The recalled ACL-QOL score differed statistically from the actual score by -3.24 points, but the authors considered this difference not clinically important. No internal consistency, structural validity, responsiveness, MIC/MCID or AUC were evaluated for ACL-QOL in this study. WOMET, IKDC, KOOS and SF-36 were also analysed in the article but are not retained in this ACL-QOL line.
Mohtadi (1998)	ACL-QOL	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	Original development and validation study of the ACL-QOL for chronic ACL deficiency. The final instrument retained 32 items across 5 domains and used a 100-mm VAS response format. Reliability was assessed in 25 patients over a 2-week interval using paired testing and average error; no significant test-retest difference was found and the overall

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																average error was 6%, but no ICC or SEM was reported. Responsiveness was assessed against a global clinical-change judgment over approximately 6 months and matched the expected direction in 21/25 patients (84%). Construct validity was supported by score range across the disease spectrum (8-99/100, including the responsiveness sample) and by known-groups discrimination: patients selected for surgery averaged 31/100, whereas those treated nonoperatively averaged 79/100. No Cronbach alpha, factor analysis, eigenvalues, RMSEA, MIC/MCID, SRM, effect size or AUC were reported.
Paradowski et al. (2013)	KOOS	Pain 0.92; Symptoms 0.95; ADL 0.94; Sport/Rec 0.96; QOL 0.95	Pain 0.86; Symptoms 0.90; ADL 0.89; Sport/Rec 0.93; QOL 0.90	Pain 5.1; Symptoms 5.2; ADL 3.9; Sport/Rec 6.8; QOL 7.3	NR	Individual: Pain 14.1; Symptoms 14.5; ADL 10.9; Sport/Rec 18.8;	NR	NR	NR	SF-36 correlations: Pain with PF 0.68, RP 0.56, BP 0.66, GH 0.38, VT 0.38, SF 0.38, RE 0.22, MH 0.17; Symptoms with PF 0.45, RP 0.29, BP 0.41, GH 0.18, VT	Pain 0.41; Symptoms 0.47; ADL 0.53; Sport/Rec 0.88; QOL 1.38	Pain 0.39; Symptoms 0.49; ADL 0.58; Sport/Rec 0.80; QOL 1.08	8-10 points cited as general KOOS MIC; ACLR-specific MIC at 1	NR	NR	Polish cross-cultural translation and validation study of KOOS following ACL reconstruction in 72 patients. The translation followed Beaton guidelines, including forward translations from English/Swedish, back-translation,

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
						QOL 20.2. Group: Pain 1.7; Symptoms 1.7; ADL 1.3; Sport/Rec 2.2; QOL 2.4				0.21, SF 0.23, RE 0.12, MH 0.08; ADL with PF 0.65, RP 0.47, BP 0.54, GH 0.38, VT 0.35, SF 0.38, RE 0.20, MH 0.12; Sport/Rec with PF 0.67, RP 0.50, BP 0.51, GH 0.33, VT 0.32, SF 0.33, RE 0.25, MH 0.10; QOL with PF 0.64, RP 0.43, BP 0.51, GH 0.39, VT 0.36, SF 0.25, RE 0.17, MH 0.14			year not determined			expert consensus and pre-test in 10 ACL-injured patients. Internal consistency was excellent for all five subscales (alpha 0.92-0.97). Test-retest reliability at 1-year follow-up was excellent (ICC 0.86-0.93; median retest interval 7 days). SEM ranged 3.9-7.3. MDC was reported at individual level (10.9-20.2) and group level (1.3-2.4); SDC was not reported as such. Construct validity was supported because all four a priori hypotheses against SF-36 were confirmed. Responsiveness from pre-ACLR to 1-year follow-up showed the expected pattern: QOL had the largest ES/SRM, followed by Sport/Rec. No factor analysis, RMSEA, MCID or AUC were reported. No preoperative floor/ceiling effects; postoperatively, ceiling effects occurred for Pain (19%) and ADL (29%).

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
Salavati et al. (2011)	KOOS	Pain 0.91; Symptoms 0.75; ADL 0.96; Sport/Rec 0.86; QOL 0.74	Pain 0.93; Symptoms 0.85; ADL 0.91; Sport/Rec 0.75; QOL 0.89	Pain 2.2; Symptoms 3.1; ADL 2.9; Sport/Rec 2.1; QOL 2.6	NR	Pain 6.1; Symptoms 8.5; ADL 8.0; Sport/Rec 5.8; QOL 7.2	NR	NR	NR	SF-36 correlations: Pain with PF 0.58, RP 0.51, BP 0.79, GH 0.46, VT 0.44, SF 0.42, RE 0.40, MH 0.49; Symptoms with PF 0.41, RP 0.40, BP 0.42, GH 0.49, VT 0.37, SF 0.42, RE 0.41, MH 0.48; ADL with PF 0.73, RP 0.51, BP 0.69, GH 0.41, VT 0.41, SF 0.53, RE 0.44, MH 0.41; Sport/Rec with PF 0.72, RP 0.49, BP 0.41, GH 0.46, VT 0.44, SF 0.46, RE 0.44, MH 0.41; QOL with PF 0.52, RP 0.54, BP 0.49, GH 0.41, VT 0.53, SF 0.59, RE 0.43, MH 0.48. PH summary: Pain 0.78; Symptoms 0.61; ADL 0.74; Sport/Rec 0.71; QOL 0.68. MH summary: Pain 0.41; Symptoms 0.39; ADL 0.44; Sport/Rec 0.42; QOL 0.43	NR	NR	NR	NR	NR	Validation study of KOOS in 57 competitive athletes after ACL reconstruction, assessed 7.6 ± 2.2 months post-surgery. No floor or ceiling effects and no missing values were observed. Test-retest interval was 6-8 days. Internal consistency was acceptable to excellent across KOOS subscales (alpha 0.74-0.96), and test-retest reliability was acceptable for all subscales (ICC 0.75-0.93). SEM ranged from 2.1 to 3.1 and MDC from 5.8 to 8.5. Dimensionality/scaling was supported by corrected item-scale correlations ranging from 0.52 to 0.85; all items exceeded 0.40, with the exception that ADL item 13 correlated equally with ADL and Symptoms. Construct validity was supported by strong correlations between theoretically related KOOS and SF-36 domains: KOOS Pain with SF-36 BP (r=0.79), KOOS ADL with SF-36 PF (r=0.73), and KOOS

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																Sport/Rec with SF-36 PF (r=0.72), with generally higher correlations with SF-36 physical than mental-health domains. No responsiveness, structural validity, MIC/MCID or AUC analyses were reported.
Comins et al. (2007)	KOOS	Original KOOS subscales: Pain 0.80; Symptoms 0.63; ADL 0.91; Sport/Rec 0.80; QOL 0.75. Modified Rasch-derived scales also reported: Pain 6-item 0.72; Symptoms 6-item 0.59; Pain+Symptoms 12-item 0.78; ADL 15-item 0.89; Mobility 7-item 0.83; Modified ADL 7-item 0.80	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	Rasch re-evaluation of 200 KOOS questionnaires completed 20 weeks after ACL reconstruction. The partial credit Rasch model was used to test whether KOOS subscales functioned as unidimensional measurement scales. Only Sport/Rec (WP chi-square=9.17, df=10, p=0.516; PSI=0.81; alpha=0.80) and QOL (WP chi-square=2.91, df=8, p=0.940; PSI=0.75; alpha=0.75) fitted the Rasch model as originally specified. The original WOMAC-derived domains did not fit: Pain (WP chi-square=36.11, df=18, p=0.007; PSI=0.81; alpha=0.80), Symptoms (WP chi-

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																square=24.65, df=14, p=0.038; PSI=0.70; alpha=0.63), and ADL (WP chi-square=67.16, df=34, p=0.001; PSI=0.87; alpha=0.91). After rescoring/item deletion, several modified scales fitted the model, but the original KOOS total structure was not supported. No ICC, SEM, SDC/MDC, external correlation coefficient, responsiveness index, MIC/MCID or AUC was reported.
Roos et al. (1998)	KOOS	NR	Pain 0.85; Symptoms 0.93; ADL 0.75; Sport/Rec 0.81; QOL 0.86	NR	NR	NR	NR	NR	NR	Key KOOS-SF-36 correlations: ADL-SF-36 physical function r=0.57; Sport/Rec-SF-36 physical function r=0.47; Pain-SF-36 bodily pain r=0.46. Full SF-36 correlation matrix reported in Table 3.	Pain 0.84; Symptoms 0.87; ADL 0.94; Sport/Rec 1.16; QOL 1.65	NR	NR	NR	NR	Original KOOS development/clinical validation study in 21 subjects undergoing ACL reconstruction. KOOS includes five separately scored domains: Pain, Symptoms, ADL, Sport/Rec, and QOL. Test-retest reliability was assessed preoperatively in 13 participants with a mean interval of 3.6 +/- 2.6 days, yielding ICCs of 0.85, 0.93, 0.75, 0.81, and 0.86 for Pain, Symptoms, ADL, Sport/Rec and QOL, respectively. Construct validity was assessed through

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																Spearman correlations with SF-36; the highest hypothesized correlations were ADL with SF-36 physical function (r=0.57), Sport/Rec with SF-36 physical function (r=0.47), and Pain with SF-36 bodily pain (r=0.46). Responsiveness at 6 months postoperatively was evaluated by effect size, with all KOOS subscales showing large effects: Pain 0.84, Symptoms 0.87, ADL 0.94, Sport/Rec 1.16, and QOL 1.65. No Cronbach alpha, SEM, SDC/MDC, factor-analytic indices, SRM, MIC/MCID or AUC were reported.
Ebrahimzadeh et al. (2014)	IKDC-SKF / IKDC Subjective Short Form	0.845; standardized-items alpha 0.915	0.845; 95% CI 0.75-0.91	NR	NR	NR	NR	NR	NR	SF-36: PF 0.522; RP 0.391; BP 0.679; GH 0.336; VT 0.402; SF 0.385; RE 0.167; MH 0.196; PCS 0.626; MCS 0.159 (p=0.055). Test-retest item/question correlation r=0.77.	NR	NR	NR	NR	NR	Persian translation/cultural adaptation and validation of the IKDC Subjective Short Form in 145 patients with ACL tear. The retest subsample included 50 randomly selected patients who completed the IKDC again after 3 days before any major treatment. Internal consistency was

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																Cronbach alpha=0.845, with standardized-items alpha=0.915. Test-retest reliability was ICC=0.845 (95% CI 0.75-0.91), and the correlation between Persian IKDC questions across test-retest was r=0.77 (p=0.000). Construct validity was tested against the Persian SF-36: PF r=0.522, RP r=0.391, BP r=0.679, GH r=0.336, VT r=0.402, SF r=0.385, RE r=0.167, MH r=0.196, PCS r=0.626 and MCS r=0.159 (MCS p=0.055). Mean total IKDC score was 39.9 +/- 18.1, with summed scores ranging from 24 to 92. SEM, SDC/MDC, structural validity indices, responsiveness metrics, MIC/MCID and AUC were not reported.
Ra et al. (2014)	IKDC-SKF / IKDC subjective score	NR	NR	NR	NR	NR	NR	NR	NR	LSI correlation: all subjects r=0.492 (95% CI 0.34-0.62) at 6 mo; r=0.296 (95% CI 0.12-0.46) at 12 mo. Ordinary recreational	NR	NR	NR	NR	NR	Comparative ceiling-effect and concurrent-validity study after ACL reconstruction (N=134; ordinary recreational subgroup n=100; elite athletes n=34). IKDC subjective score:

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
										group: r=0.452 (95% CI 0.26-0.61) at 6 mo; r=0.262 (95% CI 0.06-0.45) at 12 mo.						mean(SD) 84.5(12.1) at 6 mo and 89.8(9.9) at 12 mo. Ceiling effect: all subjects 5.2% (7/134) at 6 mo and 17.2% (23/134) at 12 mo; ordinary recreational group 1.0% (1/100) and 14.0% (14/100); elite athletes 17.6% (6/34) and 26.5% (9/34). Kolmogorov-Smirnov normality: accepted at 6 mo (p=0.0530), rejected at 12 mo (p=0.0001). Correlations with one-leg hop test LSI were weak-to-moderate and statistically significant. No alpha, ICC, SEM/SDC/MDC, structural validity, responsiveness, MIC/MCID or AUC reported.
Ra et al. (2014)	Lysholm Knee Score	NR	NR	NR	NR	NR	NR	NR	NR	LSI correlation: all subjects r=0.355 (95% CI 0.18-0.51) at 6 mo; r=0.241 (95% CI 0.06-0.41) at 12 mo. Ordinary recreational group: r=0.391 (95% CI 0.19-0.56) at 6 mo; r=0.192 (95% CI -0.02 to 0.39) at 12 mo.	NR	NR	NR	NR	NR	Comparative ceiling-effect and concurrent-validity study after ACL reconstruction (N=134; ordinary recreational subgroup n=100; elite athletes n=34). Lysholm score: mean(SD) 90.4(9.1) at 6 mo and 93.1(7.4) at 12 mo. Ceiling effect: all subjects 14.9% (20/134) at 6 mo and 30.6% (41/134) at 12

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																mo; ordinary recreational group 12.0% (12/100) and 29.0% (29/100); elite athletes 23.5% (8/34) and 35.3% (12/34). Kolmogorov-Smirnov normality rejected at both 6 mo (p=0.004) and 12 mo (p=0.0005), with negative skewness. Correlations with one-leg hop test LSI were weak-to-moderate; ordinary recreational group correlation at 12 mo was not statistically significant. No alpha, ICC, SEM/SDC/MDC, structural validity, responsiveness, MIC/MCID or AUC reported.
Çelik et al. (2014)	IKDC-SKF / IKDC Subjective Knee Form	0.89	0.91	6.0	NR	16.4	NR	NR	NR	Lysholm r=0.64; Kujala AKPS r=0.89; SF-36 PF r=0.69; RP r=0.53; BP r=0.47; GH r=0.32; VT r=0.24; SF r=0.40; RE r=0.22; MH r=0.13; PCS r=0.70; MCS r=0.05	2.09; 95% CI 1.61-2.59	NR	NR	NR	NR	Turkish translation/cross-cultural adaptation and validation of IKDC Subjective Knee Form. Baseline validation n=103; test-retest n=58, interval 5.6 +/- 2.2 days; responsiveness n=33 after ACL reconstruction. Mean IKDC preop 48.8 +/- 14.9 and 1-year postop 80.1 +/- 17.9. SEM=6.0;

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																MDC95=16.4; no floor/ceiling effects.
van Meer et al. (2013)	KOOS	NR	Pain 0.87; Symptoms 0.81; ADL 0.85; Sport/Rec 0.81; QOL 0.83	Pain 6.6; Symptoms 9.1; ADL 7.8; Sport/Rec 12.7; QOL 7.6	Individual: Pain 18.3; Symptoms 25.2; ADL 21.6; Sport/Rec 35.2; QOL 21.1	NR	NR	NR	NR	Construct validity correlations: Pain with VAS pain 0.66, SF-36 PF 0.54, SF-36 BP 0.62, Lysholm 0.68; Symptoms with VAS 0.59, SF-36 PF 0.38, BP 0.49, Lysholm 0.65; ADL with VAS 0.58, SF-36 PF 0.55, BP 0.56, Lysholm 0.71; Sport/Rec with VAS 0.47, SF-36 PF 0.62, BP 0.57, Lysholm 0.61; QOL with VAS 0.29, SF-36 PF 0.41, BP 0.36, Lysholm 0.36. Responsiveness/ change-score correlations: Pain ΔVAS 0.49, ΔSF-36 PF 0.58, ΔBP 0.55, ΔLysholm 0.55, PRI instability 0.04; Symptoms 0.35, 0.42, 0.31, 0.51, 0.15; ADL 0.44, 0.57, 0.42, 0.50, 0.08; Sport/Rec 0.55, 0.61, 0.49, 0.64, 0.14; QOL 0.16,	Pain 0.60; Symptoms 0.55; ADL 0.58; Sport/Rec 0.77; QOL 1.51	NR	NR	NR	NR	Comparative validation study of KOOS and IKDC subjective in recent ACL rupture/first year after ACL reconstruction. KOOS content validity: Pain 5/9 relevant items, Symptoms 6/7, ADL 5/17, Sport/Rec 5/5, QOL 4/4. Baseline ceiling effects: Pain 20%, ADL 46%, Sport/Rec 5%, no ceiling for Symptoms/QOL; floor effect only Sport/Rec 11%. Construct validity hypotheses confirmed: Pain 13/19 (68%), Symptoms 14/19 (74%), ADL 10/19 (53%), Sport/Rec 15/19 (79%), QOL 13/19 (68%). Responsiveness hypotheses confirmed: Pain 12/21 (57%), Symptoms 14/21 (67%), ADL 10/21 (48%), Sport/Rec 10/21 (48%), QOL 14/21 (67%). Article concludes KOOS is less useful than IKDC for short-term monitoring after ACL rupture/reconstruction.

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
										0.37, 0.14, 0.39, 0.40.						
van Meer et al. (2013)	IKDC-SKF / IKDC subjective score	NR	0.93; 95% CI 0.89-0.96	4.4	Individual SDC=12.2	NR	NR	NR	NR	Construct validity correlations: VAS pain 0.48; SF-36 physical functioning 0.67; SF-36 bodily pain 0.65; Lysholm 0.62; 16/19 construct-validity hypotheses confirmed (84%). Responsiveness/change-score correlations: ΔVAS pain 0.49; ΔSF-36 physical functioning 0.57; ΔSF-36 bodily pain 0.50; ΔLysholm 0.47; PRI instability 0.11; 18/21 responsiveness hypotheses confirmed (86%).	1.36	NR	NR	NR	NR	Comparative validation study of KOOS and IKDC subjective in recent ACL rupture/first year after ACL reconstruction. IKDC content validity: 16/18 items relevant (89%); no baseline floor or ceiling effects; ICC=0.93, SEM=4.4, individual SDC=12.2. Construct validity acceptable with 16/19 hypotheses confirmed (84%). Responsiveness acceptable with 18/21 hypotheses confirmed (86%) and effect size 1.36. Article concludes IKDC subjective is more useful than KOOS for short-term monitoring after ACL rupture/reconstruction.
Padua et al. (2004)	IKDC-SKF / IKDC Subjective Knee Form	0.91	0.90; p<0.001	NR	NR	NR	NR	NR	NR	SF-36: PF 0.67; RP 0.56; BP 0.75; GH 0.26 (NS); VT 0.36; SF 0.58; RE 0.44; MH 0.65; PCS 0.60; MCS 0.40	NR	NR	NR	NR	NR	Italian cross-cultural adaptation and validation of the IKDC Subjective Knee Form in 50 patients undergoing ACL reconstruction, with a retest

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
																subsample of 20 patients over a 5-day interval. IKDC score: mean 59.38, SD 22.88, median 55.74, range 19-97. Test-retest reliability was ICC=0.90 (p<0.001), and internal consistency was Cronbach alpha=0.91. No maximum or minimum IKDC scores were observed, indicating no floor/ceiling effects. Construct validity was assessed against SF-36 domains/components: PF 0.67, RP 0.56, BP 0.75, GH 0.26 (NS), VT 0.36, SF 0.58, RE 0.44, MH 0.65, PCS 0.60, MCS 0.40. SEM, SDC/MDC, structural validity indices, responsiveness metrics, MIC/MCID and AUC were not reported.
Irrgang et al. (2001)	IKDC-SKF / IKDC Subjective Knee Form	0.92	0.94; 95% CI 0.88-0.97; abstract reports 0.95	4.6 (test-retest); 5.4 alpha-based	NR	9.0 points (true-change threshold)	9.03; next components 1.76 and 1.19	50.2% first component; 9.8% second component	NR	SF-36: physical function 0.63; role-physical 0.47; bodily pain 0.64; general health 0.30; vitality 0.39; social function 0.47; role-emotional 0.26; mental	NR	NR	NR	NR	NR	Original development and validation study of the IKDC Subjective Knee Form. Final validation sample included 533 patients with varied knee conditions; test-retest sample n=33. PCA/factor analysis showed one dominant

Study	PROM	Cronbach's alpha	ICC	SEM	SDC	MDC	Eigenvalue	% variance	RMSEA	Correlation r	Effect size	SRM	MIC	MCID	AUC	Key extracted metrics
										health 0.25; PCS 0.66; MCS 0.16						component (eigenvalue 9.03 vs 1.76 and 1.19), explaining 50.2% of variance, supporting a single total score. Internal consistency was Cronbach alpha=0.92. Test-retest reliability was ICC=0.94 (95% CI 0.88-0.97; abstract rounds/reports 0.95), with SEM=4.6 and a 9.0-point threshold for true change; alpha-based SEM was 5.4 with 95% CI=10.5. Construct validity was supported by stronger correlations with SF-36 physical domains/components (PF 0.63, BP 0.64, PCS 0.66) than mental domains/components (RE 0.26, MH 0.25, MCS 0.16). No floor effect and minimal ceiling effect were reported (0.2% at 100). Responsiveness, MIC/MCID and AUC were not assessed in this paper.

Note 4. Values are reported exactly as extracted from each study whenever available. NR indicates that the metric was not reported in the article; NA indicates that the metric was not applicable to the instrument or analysis, for example internal consistency for a single-item scale. When an article reported MDC but not SDC, the value was retained under MDC and SDC was coded NR; conversely, when SDC was reported but MDC

was not, MDC was coded NR. For multi-domain PROMs, values are presented by domain/subscale whenever the article did not provide a total score.

Abbreviations Supplementary Table 5: ACL, anterior cruciate ligament; ACL-QOL, Anterior Cruciate Ligament Quality of Life questionnaire; ACL-RSI, Anterior Cruciate Ligament Return to Sport after Injury scale; ACLR, anterior cruciate ligament reconstruction; ADL, activities of daily living; AKPS, Anterior Knee Pain Scale/Kujala Patellofemoral Score; AUC, area under the curve; BP, bodily pain; CI, confidence interval; CKRS, Cincinnati Knee Rating System; ES, effect size; FIQ, Functional Index Questionnaire; GH, general health; GPE, Global Perceived Effect; HSS Pedi-FABS/Pedi-FABS, Hospital for Special Surgery Pediatric Functional Activity Brief Scale; ICC, intraclass correlation coefficient; IKDC, International Knee Documentation Committee; IKDC-SKF, IKDC Subjective Knee Form; I-PRRS, Injury-Psychological Readiness to Return to Sport scale; K-SES, Knee Self-Efficacy Scale; KOOS, Knee injury and Osteoarthritis Outcome Score; KOOS-Child, Knee injury and Osteoarthritis Outcome Score for Children; LSI, limb symmetry index; MCS, mental component summary; MCID, minimal clinically important difference; MDC, minimal detectable change; MH, mental health; MHLC-C, Multidimensional Health Locus of Control, chance subscale; MIC, minimal important change; NA, not applicable; NPRS, Numeric Pain Rating Scale; NR, not reported; PCS, physical component summary; Pedi-IKDC, Pediatric International Knee Documentation Committee Subjective Knee Form; PF, physical functioning; PFPS, patellofemoral pain syndrome; PROM, patient-reported outcome measure; PSS, Pain Severity Scale; QoL/QOL, quality of life; RE, role emotional; RMSEA, root mean square error of approximation; RP, role physical; SDC, smallest detectable change; SEM, standard error of measurement; SF, social functioning; SF-36, 36-Item Short Form Health Survey; Sport/Rec, sport and recreation function; SRM, standardized response mean; TSK, Tampa Scale for Kinesiophobia; VAS, visual analogue scale; VT, vitality; WOMET, Western Ontario Meniscal Evaluation Tool.