Annals of Physical and Rehabilitation Medicine 65 (2022) 101608

Available online at ScienceDirect www.sciencedirect.com

Elsevier Masson France EM|consulte www.em-consulte.com

Review

Assessing knee functionality: Systematic review of validated outcome measures

Gema Chamorro-Moriana[a,*], Veronica Perez-Cabezas[b], Fernando Espuny-Ruiz[a], Dolores Torres-Enamorado[c], Carmen Ridao-Fernández[a]

[a] Department of Physiotherapy, University of Seville, Seville, Spain
[b] Department of Nursing and Physiotherapy, University of Cadiz, Cadiz, Spain
[c] Department of Nursing, University San Juan de Dios (Bormujos) Seville, Spain

ARTICLE INFO

Article History:
Received 11 November 2020
Accepted 8 October 2021

Keywords:
Knee function
Knee injuries
Scale
Validation study
Reliability
Diagnosis

ABSTRACT

Background: Functional rating scales allow clinicians to document and quantify alterations and progression of recovery processes. There is neither awareness of numerous knee scales nor are they easy to find or compare to select the most suitable.
Objectives: We aimed to compile validated knee functional rating tools and analyze the methodological quality of their validation studies. Also, we aimed to provide an operational document of the outcome measures addressing descriptions of parameters, implementations, instructions, interpretations and languages, to identify the most appropriate for future interventions.
Methods: A systematic review involved a search of PubMed, Web of Science, CINAHL, Scopus, and Dialnet databases from inception through September 2020. The main inclusion criteria were available functional rating scales/questionnaires/indexes for knees and validation studies. Methodological quality was analyzed with the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) and COnsensus-based Standards for the selection of health Measurement Instruments Risk of Bias (COSMIN-RB).
Results: We selected 73 studies. The studies investigated 41 knee rating tools (general, 46%, and specific, 54%) and 71 validations, including 29,742 individuals with knee disorders. QUADAS-2 obtained the best results in *patient selection* and *index test* (*applicability* section). COSMIN-RB showed the highest quality in *construct validity* (most analyzed metric property). The specific tools were mainly designed for prosthesis and patellofemoral and anterior cruciate ligament injuries. More considered issues were specific function (93%), especially gait, pain/sensitivity (81%), and physical activity/sports (56%).
Conclusions and implications: We conducted a necessary, useful, unlimited-by-time and feasible compilation of validated tools for assessing knee functional recovery. The methodological quality of the validations was limited. The best validations were for the Copenhagen Knee Range of Motion Scale in osteoarthritis and arthroplasties, Knee Outcome Survey Activities of Daily Living and Lysholm Knee Score for general knee disorders and the Tegner Activity Score for anterior cruciate ligament injuries. The operational document for the scales provides necessary data to identify the most appropriate.
© 2021 The Author(s). Published by Elsevier Masson SAS. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

1. Introduction

Assessment methods for musculoskeletal disorders are necessary to identify structural, biomechanical and functional limitations, develop treatment plans and assess the effect of treatments [1]. This is why numerous efficient, valid and reliable evaluation procedures have been designed.

Knee recovery constitutes an important part of treatment for trauma in patients, such as those with osteoarthritis [2], fractures [3,4], etc., or those who undergo replacement [5,6], menisectomy [7,8] or ligament reconstructive surgery [9,10], among others. Evaluation methods such as functional rating scales allow clinicians

Abbreviations: ACL, anterior cruciate ligament; ACL-RSI, ACL-Return to Sport after Injury questionnaire; BADL, basic activities of daily living; CKRS, Copenhagen Knee Range of Motion Scale; COSMIN-RB, COnsensus-based Standards for the selection of health Measurement Instruments Risk of Bias checklist; KOOS, Knee injury and Osteoarthritis Outcome Score; KOOS-Child, KOOS for Children; KOS-ADL, Knee Outcome Survey Activities of Daily Living scale; KQOL-26, Knee Quality Of Life 26-item questionnaire; KSS, Knee Society clinical rating System; MKQ, Munich Knee Questionnaire; New-KSS, New Knee Society Scoring System; PEDI-IKDC, Pedi-International Knee Documentation Committee subjective knee form; PRISMA, Preferred Reporting Items for Systematic reviews and Meta-Analyses; QoL, quality of life; QUADAS-2, Quality Assessment of Diagnostic Accuracy Studies; ROM, range of motion; TAS, Tegner Activity Score; VISA-P, Victorian Institute of Sport Assessment for Patella; WOMET, Western Ontario Meniscal Evaluation Tool

* Corresponding author at: C/ Avicena s/n. C.P. 41009, Sevilla, Spain.
E-mail address: gchamorro@us.es (G. Chamorro-Moriana).

https://doi.org/10.1016/j.rehab.2021.101608
1877-0657/© 2021 The Author(s). Published by Elsevier Masson SAS. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

(physicians, surgeons, physiotherapists) to examine, document, describe and quantify alterations and the progression of patients during functional recovery [2,4]. Consequently, these tools promote rigorous clinical judgment, intervention protocols and effective treatments [11].

The tendency toward objectivity of these processes [1,12] has led to the development of a variety of assessment scales/questionnaires/indexes that are more viable than instrumented analysis. Such tools contribute to advantages such as easier and faster use [3], low cost and no need for specific equipment or spaces [2,13]. Thus, they are effective clinical alternatives, probably becoming the most commonly used methods [14]. However, despite many scales in knee orthopedics, many are unknown and are difficult to find or compare to select the most appropriate one in each case [15]. The various indications, heterogeneous items or components, and different rules of use and interpretations should be considered before selecting a rating scale. There are useful literature reviews of scales addressing a particular population, such as a pathology [16]. However, many of the specific scales have been validated and then applied in other contexts, thus becoming more general. Hence, clinicians should consider them all. Moreover, knee injuries are often complex (i.e., meniscopathies, anterior cruciate ligament [ACL] injuries and instability). Therefore, in addition to a compilation of tools by pathology, a global review is useful.

This study aimed to compile validated functional knee rating tools and analyze the methodological quality of their validation studies. Also, we aimed to provide an operational document of the outcome measures by addressing descriptions of parameters, implementations, instructions, interpretations and languages of the scales to identify the most appropriate for use in future interventions. The hypothesis of this study is that there is evidence on numerous and heterogeneous validated knee scales in terms of methodological quality, specificity of the target populations, content or application, among other characteristics.

2. Methods

This systematic review was based on the PRISMA guidelines [17].

2.1. Data sources and search strategy

We conducted an electronic search in PubMed, Web of Science, CINAHL, Scopus and Dialnet databases from inception through September 30, 2020. The reference lists of selected articles and the one for this review were screened for studies of interest. MeSH terms for the search strategy were divided into 2 groups: Identifier 1: “scale”, “scor*” (score/scores/scoring/scored), “questionnaire”, “test”, “index”, “assess*” (assess/assesses/assessed/assessing/assessment/assessments), “examination”, “measure”, “evaluation” and “rating”; and Identifier 2: “knee”. The search strategy was Identifier 1 AND Identifier 2 for each database [i.e., (scale OR scor* OR questionnaire OR test OR index OR assess* OR examination OR measure OR evaluation OR rating) AND knee]. All database searches were filtered by language: English, Spanish and French. The PubMed search was filtered by humans.

2.2. Study selection and inclusion criteria

The included papers met the following criteria: studies describing original validated tools based on a conceptual framework created to assess knees reported in any language or original validation studies and other subsequent ones, including physical tools or not; studies involving humans; and reports in English, Spanish and French. Tools designed to evaluate lower limbs in general were excluded.

The reviewers FE/VE and CR/DT separately screened titles and abstracts of the search results to check if the studies met the inclusion criteria. After removing duplicates, FE and CR screened the first half of the records and VE and DT the second half. GC resolved any disagreements. The full texts of the appropriate studies were obtained, and the causes for any exclusion were documented at this stage.

2.3. Data extraction

Data extraction was carried out by one reviewer (GC) and checked for accuracy by 2 more reviewers: FE/DT for tools and CR/DT for validation quality (CR the first half and DT the second half). Disagreements between reviewers were resolved by consensus or by a fourth reviewer (CR for tools; VP for validation quality). The reviewers were not blinded to authors, date of publication or journal.

The first pre-established descriptive table synthesized evidence related to functional assessment scales/questionnaires/indexes, such as sources, original and complementary validation studies, indications/implementations, countries of origin, languages, descriptions, operating instructions and observations. A second pre-established table included information on scale items and components: range of motion (ROM), strength/fatigue, instability, deformity (e.g., alignment, aesthetics), clinical signs (effusion, crepitus, locking, stiffness, etc.), diagnostic tests (radiography, nuclear magnetic resonance), physical examination tests (e.g., Lachman, Daniel, pivot shift, anterior/posterior drawer), pain and sensitivity (e.g., hypo/hyperesthesia, paresthesia), specific functions (gait, balance, stairs, squatting, etc.), quality of life (QoL) (sleep, chores, driving, leisure, etc.), psychological aspects (confidence, frustration, fear, etc.), professional and family areas, and physical and sport activities.

2.4. Quality appraisal

Two assessment procedures were used to analyze the methodological quality of the original and complementary validations associated with the identified scales/questionnaires/indexes: Quality Assessment of Diagnostic Accuracy Studies [18] scale (QUADAS-2) and COnsensus-based Standards for the selection of health Measurement Instruments Risk of Bias [19] checklist (COSMIN-RB).

The QUADAS-2, which analyses validation studies of diagnostic criteria, was used to evaluate *risk of bias*, involving patient selection, index test, reference standard and flow and timing; and *applicability*, involving just the 3 first items. These items determine whether there is “low”, “high” or “unclear” risk in each domain or concerns regarding applicability. The reviewers extracted the data by performing a general reading of the article and then a more selective and in-depth reading. The scale was completed according to these data. The Agency for Healthcare Research and Quality, Cochrane Collaboration and UK National Institute for Health and Clinical Excellence recommend QUADAS-2 for use in systematic reviews related to diagnostic accuracy studies [18].

The COSMIN-RB classifies each assessment as “very good”, “adequate”, “doubtful” and “inadequate”. It assesses the metric properties of the studies in 10 sections. The first is for the development of the tool, evaluating its design and developing a cognitive interview or pilot test [19]. The other 9 sections are for content validity, structural validity, internal consistency, reliability, measurement error, criterion validity, construct validity, responsiveness and cross-cultural validity (excluded from this study) [20]. COSMIN-RB considers only those metric properties developed in papers [19].

3. Results

3.1. Search results

The search produced 346,299 records. After the removal of duplicates, 170,911 articles were screened by title, abstract and full-text according to the inclusion criteria. Then 73 records remained for inclusion: 41 tools and 64 validation studies reporting 71 original and complementary validations. Of note, some 64 validation studies

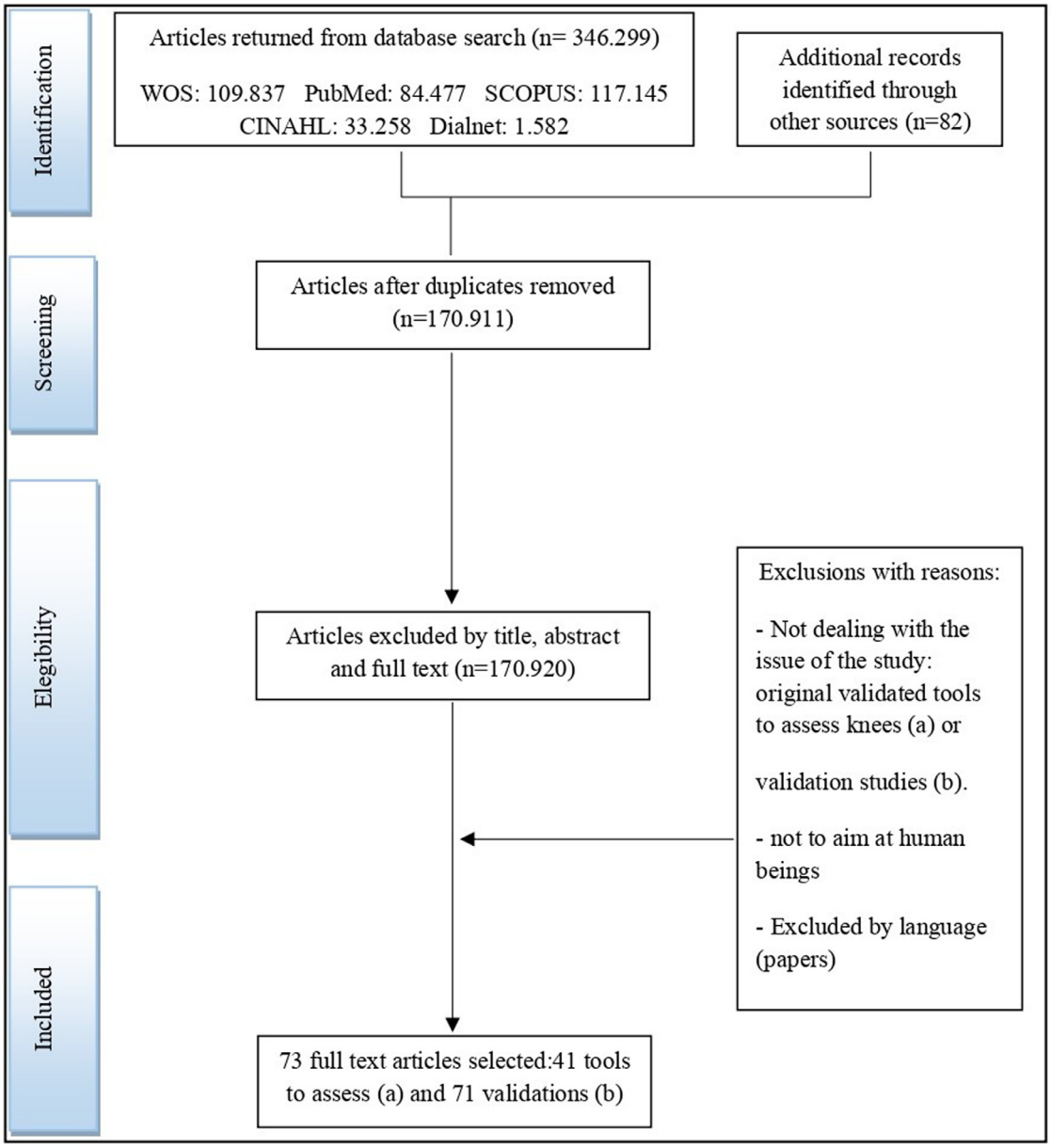

Fig. 1. PRISMA flow diagram for selection of studies.

included more than one validation, and 9 of 41 original validated tools were found in other articles. Fig. 1 shows the search and study selection process based on PRISMA guidelines [17].

3.2. Characteristics of the included tools

A detailed list and descriptive summary of the functional assessment scales/questionnaires/index characteristics, such as sources (authors and design years), original validation studies, other complementary validations, descriptions (e.g., components), operating instructions and observations (e.g., recommendations, references of physical scales), and specific references are in Table 1. This table is complemented with indications/implementations, countries of origin, languages and references in Appendix A.

3.3. Assessment of quality

Table 1 shows the original validation articles of the 41 scales. However, many were validated again later. These data were complemented by QUADAS-2 [18] and COSMIN-RB [19] evaluations including information on the methodological quality of the original and complementary validation studies selected.

QUADAS-2: *applicability* obtained better results than *risk of bias* (Appendix B). In summary, the results were as follows. *Applicability*: 188/213 (88%) positive outcomes favourably highlighting *patient selection* and *index test*, with 69/71 (97%) and 70/71 (99%) "low" concerns, respectively. *Risk of bias*: 54/284 (19%) at "low" risk and 27/284 (10%) "high" risk. The remaining assessments were unclear or could not be achieved, so they were "not applicable". The item *flow and timing* was the worst evaluated, with 15/71 showing "high" risk.

The results of COSMIN-RB are shown in Appendix C. The following results are notable.

C*onstruct validity* was the most evaluated property: 57/71 (80%) validations. This evaluation obtained the highest number of positive results [41/57 (72%) "very good" or "adequate"] and the lowest number of negative results [5/57 (9%) "inadequate"].

Patient-Reported Outcome Measure Development was the worst evaluated item, with 27/30 (90%) "inadequate" evaluations.

Table 1
Source and characteristics of included scales.

Tools. Authors, year of publication	Original validation studies other subsequent validation studies	Descriptions and operating instructions observations (recommendations, physical scales, etc.)
1. ACL-Return to sport after injury questionnaire (ACL-RSI). (Webster et al. 2008) [10]	Webster et al. 2008 [10].	*Self-administered.* Evaluation: time required to return to sports after ACL injury, and psychological factors associated with sports practice afterwards. 12 items: emotional aspect (5), confidence in own performance (5) and risk assessment (2). Items are scored from 0 to −10 following a visual analogue scale. 0 points=great disability; 120points = optimal state.
2. Activity rating scale (ARS) (Marx et al. 2001) [3]	Marx et al. 2001 [3].	*Self-administered.* Evaluation: participation in various kinds of sports according to physical activity and times the subject practiced it in the last year in the highest performance period. 4 items: run, jump, deceleration and change of direction. Items are scored: from 0- to 4, 0 being less than once a month and 4 performing the activity 4 or more times in a week. Total score ranges from 0 to 16, with 0 indicating the most deficient. Specific recommendations: to be used together with other scales.
3. British orthopedic association knee function assessment chart (Aichrot et al. 1978) [21]	Liow et al. 2003 [22].	13 items grouped into 7 components: patient satisfaction (8 points), pain (4 points), gait (13 points), ROM (13 points), varus and valgus angulations (9 points), ability for sitting (4 points) and for climbing stairs (4 points). Items are scored from 0 to 4/5. Total score ranges from 0 to 55.
4. CIncinnati knee rating system score (Noyes et al. 1983) [4]	Barber-Westin et al. 1999 [23]. Marx et al. 2001 [24].	6 subscales: subjective assessment (20 points), activity level (15 points), physical exam (25 points), stability (20 points), radiographs (10 points) and functional test (10 points). Each item is valued differently. Scale interpretation: excellent (all subscales are scored as excellent with the exception of one of them, scored as good), good (all subscales are scored as excellent or good), passable (no subscale is scored as passable) and poor (no subscale is scored as poor). Physical scale available in Berber-Westin paper [23].
5. Copenhagen knee range of motion scale (CKRS) (Mørup-petersen et al. 2018) [25]	Mørup-Petersen et al. 2018 [25].	*Self-administered.* CKRS evaluates passively the flexion and extension range of motion of the knee with illustrations. 2 items with 11 illustrations. Total score range from 0 to 11: item 1 (0–6 points with 6 illustrations), item 2 (0–5 points with 5 illustrations). Higher values mean higher range of motion. It includes illustrations
6. Fulkerson knee instability scale (Fulkerson, 1990) [26]	Paxton et al. 2003 [27].	7items: limp, need for support, stair climbing, squatting, instability, pain, and swelling. Items are scored with variable points ranging from 0 to 45. The most heavily weighted item is pain with a maximum score of 45. This questionnaire has shown to be a useful tool in patellofemoral problems and knee ligament instability. [27,28] It is an evolution of the Lysholm questionnaire to evaluate patellofemoral symptoms and results of anterior tibial tubercle transfer [28].
7. High-flexion knee scoring (Na et al. 2012) [15]	Na et al. 2012 [15].	*Self-administered.* Evaluation: knee with innovative prosthesis, "High-Flexion", that allows maximum flexion. 9 items divided into 2 subscales: pain (2) and functionality (7). Items are scored from 1 to 5. Total score varies between 9 (marked disability) and 45 (optimal state of the subject).
8. Hospital for special surgery knee score (HSS) (Ranawat et al. 1973) [29]	Gore et al. 1986 [30].	6 components: pain (30 points), function (22 points), range of motion (18 points), muscle strength (10 points), flexion deformity (10 points) and instability (10 points). Total score range from 0 to 100 (85–100=excellent, 84–70=good, 69–60=regular, <60=poor. Points are subtracted walking aids are used or in extension deficit or varus/valgus deformity. 2 disadvantages: The score does not separate knee function and global function, and when functional parameters are included, the score decreases whether subjects are old or unhealthy.
9. Hospital for special surgery patella scale (Baldini et al. 2006) [5]	Baldini et al. 2006 [5].	5 items divided into 2 subscales: subjective assessment (2 items, 65 points) and objective assessment (3 items, 35 points). Total score ranges from 0 to 100, with the highest mark indicates the optimal state, and the lowest indicate a great disability.
10. Hughston subjective knee visual analog scale system (Flandry et al. 1991) [31]	Flandry et al. 1991 [31].	*Self-administered.* 28 items divided into 5 components: pain (4), swelling, block and stiffness (6), step (1), sports participation (5) and activities of daily living (12 items). Items are valued from 10 to 0 following a visual analog scale with 10 boxes. Total score: 0–280 (0=greater disability; 280=optimal state). Percentage calculation: (Total score/280)x100. Physical scale available in Hooper [32].
11. International knee documentation committee's (ikdc) knee ligament standard evaluation form (Hefti et al. 1993) [33]	Paxton et al. 2003 [27].	8 subscales: subjective evaluation (2 items), symptoms (4 items), ROM (2 items), stability (6 items), crepitus (3 items), harvest site pathology (1 item), radiography (3 items) and functionality (1 item). The 4 first subscales are considered for the total score and the rest offer qualitative data. Subscales are assessed with their worst item evaluation: A (normal), B (almost normal), C (abnormal) and D (very abnormal). Total score corresponds to the worst assessment obtained in the subscales.
12. International knee documentation committee's (ikdc) subjective knee form (Irrgang et al. 2001) [34]	Irrgang et al. 2001 [34]. Higgins et al. 2007 [35], Williams et al. 2020 [36].	19 items divided into 3 components: symptoms (7), sport (10) and function (2) Items are scored from 0 to 10 or from 0 to 4. Total score range from 0 to 87. Percentage calculation: (Total score/maximum score)x100. 100=optimal state of the knee, 0=great disability. 90% of items have to be answered to calculate the total score.
13. Japanese knee osteoarthritis measure (JKOM) (Akai et al. 2005) [2]	Akai et al. 2005 [2].	*Self-administered.* 25 items divided into 4 subscales: pain and stiffness (8), activities of daily living (10), general activities (5) and health condition (2). The first and second subscales are scored from 0 (nothing) to 4 (extreme). The third and fourth subscales are evaluated from 0 to 4 with different possible answers. Total score ranges from 0 to 92, where 92 indicate the highest disability.The items are aimed at lifestyle in the Japanese population.
14. Kettelkamp's knee scoring scales I (Kettelkamp et al. 1975) [37]	Kettelkamp et al. 1975 [37].	13 items divided into 2 components: pain and functionality (7) and mobility, stability and deformity (6). Items possible score vary from 0 to 26 Total score range from 0 (great disability) to 103 (optimum state).
15. Kettelkampp's knee scoring scales II (Kettelkamp et al. 1975) [37]	Kettelkamp et al. 1975 [37].	11 items divided into 2 subscales: pain and functionality (6 items, 60 points) and stability and deformity (5 items, 37 points). Items are scored in ranges that vary from 0 to 22. Total score: 0 (great disability)–97 (optimal state of knee) Knee osteotomy: good (≥75), acceptable (74–60), poor (<60). Knee arthroplasty: good (≥65), acceptable (64–55), poor (<55).
16. Knee injury and osteoarthritis outcome score (KOOS) (Roos et al. 1998) [8]	Roos et al. 1998 [8].Bekkers et al. 2009 [38], Gandek et al. 2017 [39], Goodman et al. 2020 [40].	*Self-administered.* 42 items divided into 5 subscales: pain (9 items), symptoms (7 items), activities of daily living (17 items), sport (5 items) and quality of life (QoL) (4 items). It is necessary to answer at least 5,4,9,3 and 2 items respectively in each subscale to calculate the total score. Items are scored from 0 (none) to 4 (extreme). Each subscale is analyzed independently, in a

(*continued*)

Table 1 (*Continued*)

Tools. Authors, year of publication	Original validation studies other subsequent validation studies	Descriptions and operating instructions observations (recommendations, physical scales, etc.)
		range from 0 (extreme difficult) to 100 (optimal state), using a specific formula (available at www.koos.nu). Physical scale available in the KOOS website (www.koos.nu).
17. Knee injury and osteoarthritis outcome score for children (KOOS-child) (Örtqvist et al. 2012) [41]	Örtqvist et al. 2012 [41]. Örtqvist et al. 2014 [42].	*Self-administered.* 5 subscales: pain (8 items), symptoms (7 items), activities of daily living (11 items), functionality in sports (7 items) and QoL (6 items). Items are scored from0 (nothing) to 4 (extreme), following a Likert scale. Each subscale is valued separately, and at least 4,4,6,4 and 3 items have to be answered, respectively. A specific formula (www.koos.nu) transforms the final result in a range from 0 (great disability) to 100 (optimal state). Specific recommendations**:** applicable to children between 7 and 16 years of age. Observations**:** maybe the youngest children need help to read, but they can answer themselves. Physical scale available in the KOOS website (www.koos.nu).
18. Knee injury and osteoarthritis outcome score for joint replacement (KOOS-JR) (Lyman et al. 2016) [6]	Lyman et al. 2016 [6]. Hunnicutt et al. 2019 [43], Buller et al. 2020 [44].	*Self-administered.* 7 items: rigidity, pain and activities of daily living. They are scored from 0 (nothing) to 4 (extreme), following a Likert scale. A specific formula (www.koos.nu) transforms the final result (0 to 28) in a range from 0 (great disability) to 100 (optimal state). Physical scale available in the KOOS website (www.koos.nu).
19. Knee injury and osteoarthritis outcome score for patellofemoral pain and osteoarthritis (KOOS-PF) (Crossley et al. 2018) [45]	Crossley et al. 2018 [45].	*Self-administered.* Evaluation: symptoms relevant to people with patellofemoral pain and/or OA, and change over time. 11 items divided into 3 subscales: stiffness (1), pain (9), QoL (1). Items are scored from 0 (none) to 4 (extreme) following the Likert scale. Total score is transformed into a ranging from 0 (maximum disability) to 100 (optimal state) [45] by a specific formula (www.koos.nu). Physical scale available in the KOOS website (www.koos.nu).
20. Knee injury and osteoarthritis outcome score-physical function short-form (KOOS-PS) (Perruccio et al. 2008) [46]	Perruccio et al. 2008 [46], Davis et al. 2009 [47]. Ruyssen-Witrand et al. 2011 [48].	*Self-administered.* 7 items assess the degree of difficulty in performing different activities and are scored from 0 (nothing) to 4 (extreme) following a Likert scale. Total score is transformed by a specific formula (www.koos.nu) in a range from 0 (great disability) to 100 (optimal state). Physical scale available in the KOOS website (www.koos.nu).
21. Knee injury and osteoarthritis outcome score 12 (KOOS-12) (Gandek et al. 2019) [49]	Gandek et al. 2019 [49]. Gandek et al. 2019 [50], Eckhard et al. 2020 [51].	*Self-administered.* 12 items divided into 3 subscales: pain (4), function (4) and QoL (4). Items are valued from 0 (none) to 4 (extreme). A specific formula (www.koos.nu) transforms the total result in a score from 0 (great disability) to 100 (optimal state). Physical scale available in the KOOS website (www.koos.nu).
22. Knee outcome survey activities of daily living scale (KOS-ADL) (Irrgang et al. 1998) [52]	Irrgang et al. 1998 [52]. Marx et al. 2001 [24].	*Self-administered.* 2 subscales, with 8 items each, that evaluate symptoms and functional disabilities. Items are scored from 0 (inability to perform de activity) to 5 (activity performed without difficulty), from 0 to 3 or from 0 to 2. Total score ranges from 0 to 80, while 80 indicates the optimal result and 0 a great disability of the subject.
23. Knee quality of life 26-item questionnaire (KQOL-26) (Garratt et al. 2008) [53]	Garratt et al. 2008 [53].	26 items: physical function (15), limitations in activities (5) and emotional function (6). Each one is rated between 0 and 4. Total score ranges from 0 (marked pathology conditions) to 104 (optimal state). Physical scale available in Chuang [54].
24. Knee self-efficacy scale (K-SES) (Thomeé et al. 2006) [55]	Thomeé et al. 2006 [55].	*Self-administered.* 22 items in four sections: A, daily activities (7 items); B, sports activities (5 items); C, knee function tasks (6 items), where the patients report how confident they are about performing the tasks; D, knee function in the future (4 items), where the patients report how confident they feel about their future capabilities. Items are scored in an 11-grade Likert Scale from 0 (not at all confident) to 10 (very confident). [56,57]. Patients have to read and understand Swedish. Swedish version is not available (published in English).
25. Knee society clinical rating system (KSS) (insall et al. 1989) [58]	Lingard et al. 2001 [59].	2 components: state of the knee (7 items) and functionality (3 items). The first one evaluates pain, range of motion and stability; the second analyzes gait for 100 m, going up and down stairs and the use of walking aids. Components have a maximum score of 100 (optimum state of the subject).
26. Korean knee score (KKS) (Ha et al. 2012) [60]	Ha et al. 2012 [60]. Kim et al. 2013 [61].	*Self-administered.* 41 items distributed in 4 subscales: pain and symptoms (12), functionality (17), soil-based lifestyle (6), and social-emotional function (6). Items are rated between 0 and 4. Total score is transformed into a 0 to 100-point scale (great disability-optimal state of the subject): (KKS / 164 Score) x 100. It was validated for people over 60 years.
27. Kujala patellofemoral score (KPS) (Kujala et al. 1993) [62]	Ittenbah et al. 2016 [63]. Paxton et al. 2003 [27].	*Self-administered.* 13 items, for example: limp, discharge of weight, pain and swelling, among others. Items are scored from 0 to 5 or 10 points. Total score varies between 0 (great disability) and 100(optimal knee result).
28. Lequesne algofunctional index (Lequesne et al. 1987) [64]	Lequesne et al. 1987 [64]. Faucher et al. 2002 [65].	*Self-administered.* Evaluation: severity for osteoarthritis (knee and hip).3 components: pain (5 items), maximum distance walked (1 item) and activities of daily living (4 items). Items are ranged differently with a maximum score of 24(0=no disability, 1−4=slight disability, 5−7=moderate disability, 8−10=severe disability, 11−13=very severe disability, ≥14=extremely severe disability.
29. Lysholm knee score (Tegner et al. 1985) [66]	Bengtsson et al. 1996 [67]. Briggs et al. 2006 [68] and 2009 [69], Kocher et al. 2004 [70], Marx et al. 2001 [24], and Paxton et al. 2003 [27].	*Self-administered.* 8 items, which are evaluated differently, from 0 until 25 points. The total rating can vary between 0 and 100. Interpretation: poor score (< 65 points); moderate (65−83); good (84−94); excellent (95−100).
30. Multi-ligament quality of life questionnaire (MLQOL) (Chahal et al. 2014) [71]	Chahal et al. 2014 [71].	*Self-administered.* 52 items divided into 4 components: physical impairments (19), emotional deficits (15), limitations on activity (12) and social participation (6). Each item is evaluated from 0 to 4. Item scores and are transformed into percentages (0%=optimal state of the subject, 100%=marked disability). The 4 components can be applied individually depending on the purpose of the study and/or measured aspects.
31. Munich knee questionnaire (MKQ) (Beirer et al. 2015) [72]	Beirer et al. 2015 [72].	*Self-administered.* 33 items, divided into 5 subscales: symptoms (7), pain (6), activities of daily living (5), sports activities (6) and physical function and QoL (9). Each item is valued between 0 and 10. Total score is expressed in a percentage (0%=poor result, 100%=excellent result).
32. New Knee Society Scoring System (New-KSS) (Noble et al. 2012) [73]	Noble et al. 2012 [73]. Dinjens et al. 2014 [74], Cullinton et al. 2018 [75].	34 items divided into 4 components: objective knee score (7 items, 100 points), satisfaction (5 items, 40 points), expectations (3 items, 15 points) and functional activity (19 items, 100 points). Total score ranges from 0 (great disability) to 255 (optimal state of the subject). The origin of the New-KSS is the KSS [58].

(*continued*)

Table 1 (*Continued*)

Tools. Authors, year of publication	Original validation studies other subsequent validation studies	Descriptions and operating instructions observations (recommendations, physical scales, etc.)
33. Norwich Patellar Instability (NPI) Score (Smith et al. 2014) [76]	Smith et al. 2014 [76]. Smith et al. 2019 [77]	*Self-Administered.* 19 items divided into several activities (12 high energy and 7 low energy), related to instability. They are rated on a Likert scale that varies from 0 to a maximum of 25. Total score is turned to a percentage (100%=great patellar instability).
34. Oxford Knee Score (OKS) / OXFORD 12-ITEM Knee QuestionnairE (Dawson et al. 1998) [78]	Dawson et al. 1998 [78]. Harris et al. 2013 [79], Conaghan et al. 2007 [80].	*Self-administered.* 12 items divided into 2 components: pain (5) and functionality (7). Items are rated from 1 to 5. Total score ranges from 12 to 60. > 41=excellent, 34–41=good, 27–33=fair, <27=poor.
35. Patellofemoral Pain Syndrome (PFPS) **Severity Scale** (Laprade et al. 2002) [81]	Laprade et al. 2002 [81].	*Self-administered.* 2 subscales. First subscale (8 items): climbing stairs, squatting, walking, jogging/running, participation in sports, sitting and kneeling. Second subscale: pain at rest and during rest after performing an activity. Items are scored from 0 to 10 following a visual analogue scale, and total score varies between 0 (optimal state of the subject) and 100 (great disability). If the "not attempted" box was marked because the subject did not perform the activity, it is eliminated and the final result is calculated based on 90 points instead of over 100.
36. PEDI- International knee documentation committee subjective knee form (PEDI-IKDC) (Iversen et al. 2010) [82]	Kocher et al. 2011 [83].	14 items divided into two subscales: symptoms (9 items, 49 points) and sports activities (5 items, 72 points). Items have different scores (0 to 10, 1 to 2…). Total score ranges from 18 (the highest disability) to 121 (optimal state). Equation for calculating percentages: (100 points - lowest number of points possible / range of points) x 100. It is indicated for subjects between 10 and 18 years of age.
37. Quality of life questionnaire outcome measure for chronic anterior cruciate ligament deficiency (Mohtadi 1998) [9]	Mohtadi 1998 [9].	*Self-administered.* 32 items divided into 5 subscales: physical complaints and complaints, work-related concerns, recreational and sports activities, lifestyle and emotional aspects. Each item is evaluated according to a Likert scale of 0 to 100. Total score range from 0 (the greatest disability) to 3200 (optimal state of the subject).
38. Short version of the ACL-RSI Scale (Webster et al. 2018) [84]	Webster et al. 2018 [84].	*Self-administered.* Evaluation: time required to return to sports after ACL injury, and psychological factors associated with sports practice afterwards. 6 items, graded on a scale of 0 to 100 given in 10-point increments. Total score ranges from 0 to 100, where the highest scores reflect greater psychological readiness.
39. Tegner activity score (TAS) (Tegner et al. 1985) [66]	Briggs et al. 2006 [68]. Briggs et al. 2009 [69], Paxton et al. 2003 [27].	*Self-administered.* Evaluation: work, recreational and sports activities in which the subject can participate. The score varies from 0 to 10, with 10 indicating the highest level of sport activity, and 0 a subject with a pension due to knee problems.
40. Victorian institute of sport assessment for patella (VISA-P) Score (Visentini et al. 1998) [13]	Visentini et al. 1998 [13].	*Self-administered.* 8 items,7 of them scored from 0 to 10.. The last item is assessed with 3 options (A, B and C) according to the patient's symptoms (0–30). Total score ranges from 0 to 100 (asymptomatic patient, who can perform sport without pain or restriction). In the cross-cultural adaptation into Spanish the results of some items vary with respect to the original.
41. Western ontario meniscal evaluation tool (WOMET) (Kirkley et al. 2007) [7]	Kirkley et al. 2007 [7]. Sihvonen et al. 2012 [85].	*Self-administered.* 16 items separated into 3 subscales: physical symptoms, sports activity and lifestyle, and emotions. Items are rated in a line from “no symptoms” on the left to “extreme symptoms” on the right. The clinician should measure the distance from the left side of the line and calculate the score out of 100. Total score can be calculated for each domain or added up for an aggregate score out of 1600 (the worst state of the subject).

* The first reference of each scale and its authors (first column) lead to the original validation study and to the physical tool. If these data are offered by other studies, they are referenced together with their corresponding authors in the second column (validation study) and the right column (physical tool).
** A few validation studies include more than one validation, and several original validated tools were taken from other articles.
ACL-RSI, ACL-Return to Sport after Injury questionnaire; ACL, anterior cruciate ligament; ARS, Activity Rating Scale; PCL, posterior cruciate ligament; MCL, Medial Collateral Ligament; CKRS, Copenhagen Knee Range of motion Scale; HSS, Hospital for Special Surgery knee score; IKDC, International Knee Documentation Committee's; PRP, platelet-rich plasma; ROM, range of motion; JKOM, Japanese Knee Osteoarthritis Measure; KOOS, Knee injury and Osteoarthritis Outcome Score; QoL, quality of life; KOOS-JR, Knee injury and Osteoarthritis Outcome Score for Joint Replacement; KOOS-PF, Knee injury and Osteoarthritis Outcome Score for Patellofemoral pain and osteoarthritis; KOOS-PS, Knee injury and Osteoarthritis Outcome Score - Physical function Short-form; KOS-ADL, Knee Outcome Survey Activities of Daily Living scale; KQOL-26, Knee Quality Of Life 26-item questionnaire; K-SES, Knee Self-Efficacy Scale; KSS, Knee Society clinical rating System; KKS, Korean Knee Score; KPS, Kujala Patello-Femoral Score; MLQoL, Multi-Ligament Quality Of Life questionnaire; MKQ, Munich Knee Questionnaire; NPI, Norwich Patellar Instability score; OKS, Oxford Knee Score; PFPS, Patello-Femoral Pain Syndrome severity scale; PEDI-IKDC, Pedi-International Knee Documentation Committee subjective knee form; TAS, Tegner Activity Score; VISA-P, Victorian Institute of Sport Assessment for Patella score; WOMET, Western Ontario Meniscal Evaluation Tool.

Responsiveness was assessed in 34/71 (48%) validations, with 22/34 (65%) favourable: 18/34 (53%) “very good” and 4/34 (12%) “adequate”.

Reliability was considered in 47/71 (66%) validations: 0/66 “very good” and 40/47 (86%) “inadequate”.

Internal consistency was analyzed in 39/71 (55%) validations: 12/39 (31%) “very good”, 19/39 (49%) “inadequate” and 8/39 (21%) “doubtful”.

Structural validity and *measurement error* were the least evaluated, in 7/71 (10%) and 8/71 (11%) validations, respectively. *Measurement error* was analyzed as “inadequate” in 5/8 (62%) validations.

3.4. Administration, indications and applications

In total, 30/41 (73%) scales were self-administered. Some are designed to evaluate knee dysfunctions in general (46%) and others specific disorders or contexts (54%). The latter are shown with their target population in Table 2. All tools are for adults, except for 2 [41,82], which are specific for children.

3.5. Items and components

Table 3 contributes specific data about items and components of the tools. The percentages of these areas are represented in Fig. 2. The 3 most addressed areas in the knee scales were specific function (93%), knee pain (80%) and physical activities and sport (56%).

3.6. Comparative analysis of the results

Fig. 3 shows a comparison between the methodological quality of QUADAS-2 and COSMIN-RB for each validation study, by population. In general, the comparative analysis shows that the 4 scales with the best methodological quality, according to the number of positive results by QUADAS-2 or COSMIN-RB, were self-administered general knee scales: Copenhagen Knee Range of Motion Scale (CKRS), Knee Outcome Survey Activities of Daily Living Scale (KOS-ADL) and Lysholm Knee Score (QUADAS-2), and Tegner Activity Score (TAS) (COSMIN-RB). The Lysholm Knee Score had 6 validations, and TAS, KOS-ADL and CKRS, had 3, 2 and 1, respectively. However, KOS-ADL was notable for containing more content. Kettelkamp's Knee Scoring

Table 2
Specific outcome measures and associated populations.

Populations	Scales
Arthroplasties/ replacements	High-flexion Knee Scoring [15]
	Hospital for Special Surgery patella Scale [5]
	New Knee Society Scoring System [73]
	Knee Injury and Osteoarthritis Outcome Score for Joint Replacement [6]
	Oxford Knee Score [78]
Anterior cruciate ligament injuries	Anterior Cruciate Ligament-Return to Sport after Injury (ACL-RSI) questionnaire [10]
	Knee Self-Efficacy Scale [55]
	Quality of Life Questionnaire Outcome Measure for Chronic Anterior Cruciate Ligament Deficiency [9]
	Short version of ACL-RSI [84]
Ligament injuries	International Knee Documentation Committee's knee ligament standard evaluation form [33]
	Multi-ligament Quality of Life Questionnaire [71]
Osteoarthritis	Japanese Knee Osteoarthritis Measure [2]
	Korean Knee Score [60]
	Lequesne Algofunctional Index [64]
Instability	Fulkerson Knee Instability Scale [26]
Surgery in general	Hospital for Special Surgery Knee Score [29]
Patellofemoral disorders	Kujala Patellofemoral Score [62]
	Norwich Patellar Instability [76]
	Patello-Femoral Pain Syndrome Severity Scale [81]
	Knee injury and Osteoarthritis Outcome Score for Patello-Femoral pain and osteoarthritis [45]
Patellar tendinopathy	Victorian Institute of Sport Assessment for Patella [13]
Meniscopathies	Western Ontario Meniscal Evaluation Tool [7]

ACL-RSI, Anterior Cruciate Ligament-Return To Sport After Injury.

Scale (versions I and II) and the Victorian Institute of Sport Assessment for Patella (VISA-P) were the 3 with the lowest overall methodological quality and, moreover, had only 1 validation each. Only VISA-P was self-administered and specific, for patellar tendinopathy.

In terms of population groups and following the same methodological criteria, the evaluations of the following stand out in a positive way: arthroplasties, osteotomies and osteoarthritis with the CKRS; knee disorders in general with the KOS-ADL and Lysholm Knee Score (compared above); patellofemoral disorders with the Norwich Patellar Instability; meniscopathies with the Western Ontario Meniscal Evaluation Tool (WOMET) (with 2 validations, 16 items and self-administered) and Knee Quality Of Life 26-item questionnaire (KQOL-26) (with 1 validation, 26 items but with less content and non −self-administered); and ligamentous disorders in general with the KQOL-26.

When comparing the only 2 children's scales, the Knee Injury and Osteoarthritis Outcome Score (KOOS) for Children (KOOS-Child) and Pedi-International Knee Documentation Committee subjective knee Form (PEDI-IKDC), with 2 and 1 validations, respectively, were methodologically limited in a similar way. Both scales were intended for general pathology, but the KOOS-Child covers more items and is self-administered, with pictures to facilitate completion.

4. Discussion

This systematic review compiled 41 validated knee functional assessment tools and analyzed the methodological quality of the 71 associated validations. We also provide parameter descriptions, implementations, instructions, interpretations and languages of the outcome measures to give an operative document to identify the most appropriate tool in each case.

4.1. Quality appraisal of the review

The systematic review collected reports of studies with no limit on time, so older, well-known scales and recent ones were included. In general, QUADAS-2 and COSMIN-RB showed a tendency to improve the quality of validations over the years (i.e., previously, prestigious journals demanded less scientific requirements than currently). Those are complete validations including validation of content, criteria and construct, plus reliability [12,86].

QUADAS-2 was used to analyze the *risk of bias* and concerns regarding the *applicability* of the 71 validations. Because of lack of reference standards, essential to achieve proper methodological quality [18], some items were not answered. Validations of the CKRS [25], KOS-ADL [24] and Lysholm Knee Score [24] had the best scores, with 6 of 7 positive responses. The rest were "unclear" because they did not reflect the sampling procedure, which must be consecutive or randomized for an adequate evaluation.

The COSMIN-RB was used to evaluate each metric property analysed in the validations exhaustively and individually. The validation of Briggs et al. [69] regarding TAS, obtained the best results, with 5 evaluations being "very good" or "adequate", followed by 3 positive results for the CKRS [25], KOOS [39], KOOS Joint Replacement [6], KOOS-Physical Function Short-Form [47], Munich Knee Questionnaire [72], New Knee Society Scoring System (New-KSS) [74] and WOMET [85]. Thus, the scales' methodological quality was limited. Of note, the results were conditioned by the complexity of the COSMIN-RB because each metric property is assessed with a specific question battery and marked according to the lowest result obtained. The sample size, statistical tests or sample used reduced the good results. This situation occurs in assessing *reliability*, which requires kappa indexes for ordinal scales and intraclass correlation coefficients for continuous scales, to avoid "inadequate" results. Also, essential properties for measurement instruments, such as *reliability* or *internal consistency* [87], were not present in some validations: High-flexion Knee Scoring, Hughston Subjective Knee Visual Analog Scale System, Kettelkamp's Knee Scoring Scale (versions I and II) and KSS. For these reasons, only the most rigorous validations (i.e., with high scientific quality), obtained satisfactory results by COSMIN-RB.

The data provided imply the need for improving the validation quality for future scales and complementing the existing ones, supporting tools frequently used in the clinic as shown by Briggs et al., who validated the TAS in 2006 [68] and 2009 [69], thus improving the design and statistics and, consequently, QUADAS-2 and COSMIN-RB scores.

In summary, the overall results on methodological quality showed that no scale was at high quality according to both QUADAS-2 and COSMIN-RB. Thus, the best scores for the 2 evaluations were, according to QUADAS-2, the CKRS for osteoarthritis and arthroplasties and KOS-ADL and Lysholm Knee Score for knee disorders in general, and according to COSMIN-RB, the TAS for ACL injuries. Moreover, the Lysholm Knee Score was associated with more validations, and the KOS-ADL covered more content. The 4 scales, which are self-administered, are currently used for disorders in general. Kettelkamp's Knee Scoring Scale (version I and II) and VISA-P were the worst tools according to QUADAS and COSMIN-RB.

4.2. Populations and implementations

The measurement tools were created and validated in general populations or, as in most cases, in specific ones, for clinical manifestations [81], treatments [5,29], age (e.g., infant population [41,82]), sport activities [9], lifestyles (Japanese population [2]) or pathologies [9]. Numerous tools were designed for disorders in general (i.e., KOOS [8], Lysholm Knee Score [66] and Munich Knee Questionnaire [72]) and others for particular injuries [i.e., ACL Return to Sport after

Table 3
Data approached by items and components of the scales.

Scales	Self-administered	Range of motion	Strength/ fatigue	Instability	Deformity (e.g., alingment)	Signs (e.g., effusion, crepitus, stiffness, swelling)	Diagnostic test (e.g., x-ray)	Physical examination test	Pain/ sensitivity	Knee specific function		QoL/global functions (e.g., washing, driving, leisure)	Psychological aspects (e.g., frustration, confidence, fear)	Familiar scope	Professional aspects	Physical and sport activities
										gait*/ balance	Squatting, kneeling, rising, stairs, standing, etc.					
1. Anterior Cruciate Ligament-Return to Sport after Injury questionnaire (ACL-RSI) [10]	✓												✓			✓
2. Activity Rating Scale [3]	✓										✓					✓
3. British Orthopedic Association Knee Function Assessment Chart [21]		✓			✓				✓	✓	✓		✓			
4. Cincinnati Knee Rating System Score [4]		✓		✓		✓	✓	✓	✓	✓◯	✓					✓
5. Copenhagen Knee Range of Motion Scale [25]	✓	✓														
6. Fulkerson Knee Instability Scale [26]		✓		✓	**	✓			✓	✓	✓					
7. High-flexion Knee Scoring [15]	✓								✓		✓					
8. Hospital for Special Surgery Knee Score [29]		✓	✓	✓					✓	✓	✓					
9. Hospital for Special Surgery Patella Scale [5]		✓	✓			✓			✓		✓					
10. Hughston Subjective Knee Visual Analog Scale System [31]	✓			✓		✓			✓	✓◯	✓					✓
11. International Knee Documentation Committee's knee ligament standard evaluation form [33]		✓		✓		✓	✓		✓		✓	✓				✓
12. International Knee Documentation Committee's subjective knee form [34]	✓			✓		✓			✓	✓	✓	✓				✓
13. Japanese Knee Osteoarthritis Measure [2]	✓					✓			✓	✓◯	✓	✓	✓			
14. Kettelkamp`s Knee Scoring Scale I [37]		✓			✓	✓		✓	✓	✓◯	✓	✓				
15. Kettelkamp`s Knee Scoring Scale II [37]					✓	✓		✓	✓	✓◯	✓	✓				
16. Knee injury and Osteoarthritis Outcome Score [8]	✓	✓		✓		✓			✓	✓	✓	✓	✓			✓
17. Knee injury and Osteoarthritis Outcome Score for children [41]	✓	✓				✓			✓		✓	✓	✓		✓	✓
18. Knee injury and Osteoarthritis Outcome Score for joint replacement [6]	✓	✓							✓		✓					
19. Knee injury and Osteoarthritis Outcome Score for patellofemoral pain and osteoarthritis [45]	✓					✓			✓		✓	✓				✓
20. Knee injury and Osteoarthritis Outcome Score – physical function short-form [46]									✓		✓	✓				
21. Knee injury and Osteoarthritis Outcome Score 12 [49]	✓			✓					✓	✓	✓	✓	✓			
22. Knee Outcome Survey Activities of Daily Living Scale [52]	✓	✓	✓	✓		✓			✓	✓◯	✓					
23. Knee Quality of Life 26-item questionnaire [53]	✓									✓	✓	✓	✓		✓	✓
24. Knee Self-efficacy Scale [55]	✓			✓						✓	✓	✓	✓			✓
25. Knee Society Clinical Rating System [58]		✓		✓	✓		✓		✓	✓◯	✓					
26. Korean Knee Score [60]	✓	✓				✓			✓	✓	✓	✓	✓			
27. Kujala patellofemoral score [62]	✓	✓	✓	✓	✓	✓			✓	✓	✓					✓

(*continued on next page*)

Table 3 (*Continued*)

Scales	Self-administered	Range of motion	Strength/ fatigue	Instability	Deformity (e.g., alingment)	Signs (e.g., effusion, crepitus, stiffness, swelling)	Diagnostic test (e.g., x-ray)	Physical examination test	Pain/ sensitivity	Knee specific function		QoL/global functions (e.g., washing, driving, leisure)	Psychological aspects (e.g., frustration, confidence, fear)	Familiar scope	Professional aspects	Physical and sport activities
28. Lequesne Algofunctional Index [64]	✓	✓				✓			✓	✓○	✓	✓				
29. Lysholm Knee Score [66]	✓			✓		✓			✓	✓○	✓					
30. Multi-ligament Quality of Life Questionnaire [71]	✓		✓	✓		✓			✓	✓	✓	✓	✓		✓	✓
31. Munich Knee Questionnaire [72]	✓	✓	✓	✓		✓			✓	✓	✓	✓	✓			✓
32. New knee Society Scoring System [73]		✓		✓	✓		✓		✓	✓○	✓	✓	✓			✓
33. Norwich Patellar Instability Score [76]	✓			✓						✓	✓					✓
34. Oxford Knee Score [78]	✓			✓					✓	✓○	✓	✓				
35. Patellofemoral Pain Syndrome Severity Score [81]	✓								✓	✓	✓	✓				✓
36. Pedi- International Knee Documentation Committee Subjective Knee Form [82]		✓				✓			✓	✓	✓	✓				✓
37. Quality of life Questionnaire Outcome Measure for Chronic Anterior Cruciate Ligament Deficiency [9]	✓	✓	✓	✓					✓		✓	✓	✓	✓	✓	✓
38. Short version of the ACL-RSI scale [84]	✓												✓			✓
39. Tegner Activity Score [66]	✓									✓		✓				✓
40. Victorian Institute of Sport Assessment for Patella Score [13]	✓								✓		✓					✓
41. Western Ontario Meniscal Evaluation Tool [7]	✓	✓	✓	✓		✓			✓	✓	✓		✓		✓	✓

* *o* = the scale considers orthopedic tool or person for the assisted gait.

ACL-RSI, Anterior Cruciate Ligament-Return to Sport after Injury; QoL, quality of life.

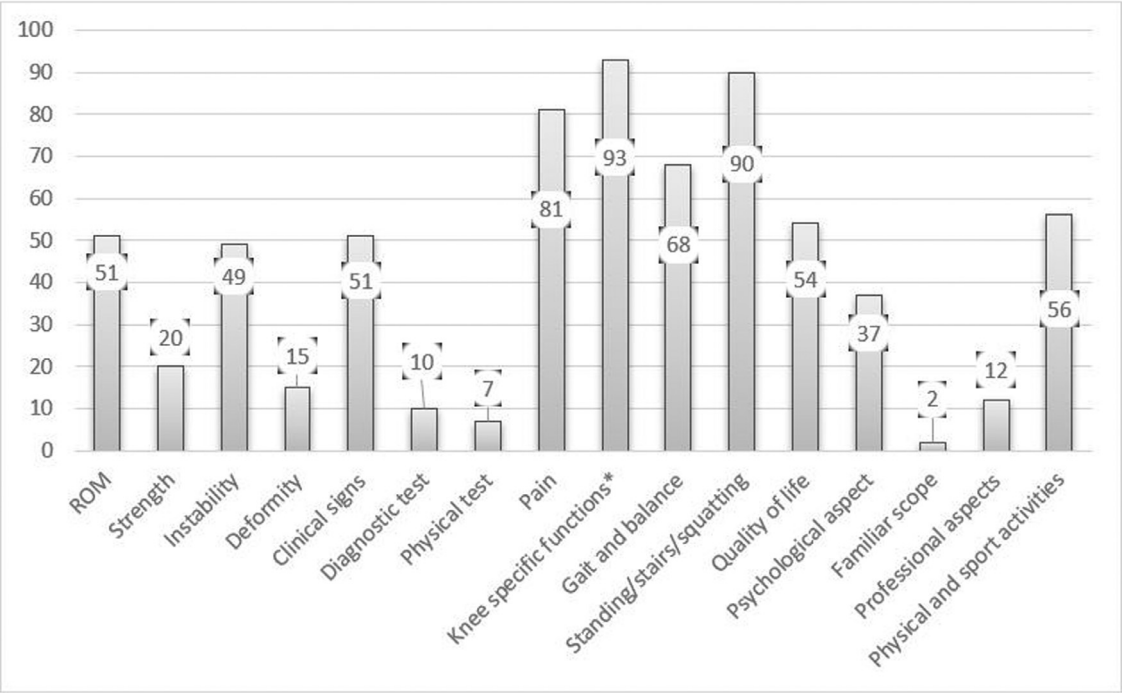

Fig. 2. Contents of the scales in percentage. * "Knee specific functions" includes the items "Gait and balance" and "Standing/stairs/squatting". ROM, range of motion.

Injury questionnaire (ACL-RSI) [9]), Japanese Knee Osteoarthritis Measure [2], and KOOS Patello-Femoral pain and osteoarthritis [45].

Therefore, despite the design of tools aimed at specific populations, some had been applied and validated subsequently in other contexts. For example, the TAS was created for ligamentous disorders [66] and subsequently validated in meniscopathies [68], and the original ACL-RSI was applied in people > 60 years old [88]. This expands the range of implementation and usefulness of the tools in clinical practice. Consequently, the initial objective may vary over time, thus resulting in a living and active scientific process, with clinicians and researchers contributing their knowledge. Other examples were the content modification of the New-KSS [73] or the adaptation to Turkish of the ACL-RSI [89]. In all cases, to ensure good clinical performance, quality cross-cultural adaptation studies will be required [90–94].

By contrast, although increasing the range of application is beneficial, in some cases such as dysfunctions or surgeries with pathognomonic profiles, specific scales would be needed to achieve more effective clinical assessments.

4.3. Items and components of outcome measures

In terms of the content of the tools, most of the scales (93%) approached some specific functions in knees, such as gait, balance, stairs, standing, squatting, sitting with knees crossed. Gait was the most-addressed component (68%), being one of the most important functions in humans [12,95,96]. Thus, it is a significant action within the basic activities of daily living (BADL) according to Barthel [97]. Gait included particularly assisted gait by someone [37] or orthopedic tools [2,31,37,52,58,64,66,73,78], time [2,37,73,78] and distance [4,37,52,58,64].

Pain was the second issue considered (81%). These are reasonable data because pain is one of the most common symptoms in knee injuries [8], for example in osteoarthritis, which affects one-third of the world population [98].

The third most approached topic was physical activity and sports (56%), which was included in the ACL-RSI [9] and KOOS [8], among others. Many knee pathologies are directly associated with certain sports. ACL and meniscal injuries represented 20% and 11% of the knee pathologies associated with sports, respectively, especially in football (35%) and skiing (26%). Lateral collateral ligament disorders are related to tennis and gymnastics, medial collateral ligament to judo and skiing, and ACL to handball and volleyball, etc. [99]

QoL was the fourth most assessed topic (54%). Knee dysfunctions negatively affect QoL. Earl-Boehm et al. [100] concluded that patello-femoral pain affected professional life and physical activity as well as BADL. Furthermore, acute pain could lead to reducing physical aspects, which affect the patient's independence and hinder psycho-emotional aspects, thus promoting conditions such as depression [101] and consequently hindering QoL.

Alterations in knee ROM (51%) lead to impairments in BADL; that is, gait is modified when there are movement limitations, especially on inclined planes or when going up or down stairs [102]. In addition, gait becomes slower and more unstable [102]. By contrast, increased knee flexion during the stance phase in osteoarthritis results in joint overload [103].

About half (51%) of the tools included clinical signs, objectively complementing the symptoms perceived by patients themselves. This review advocated for clinical signs, according to Luyten et al. [104], who addressed crepitation and radiological signs, etc.

Curiously, instability was tackled in only 49% of the scales, although it is an important concern for knee functionality [105]. Strength is even more remarkable than instability, being one of the basic physical qualities that makes up human function, together with flexibility, speed, and cardiorespiratory endurance [106]. However, it is assessed in only 20% of the scales.

Psycho-emotional aspects are considered in 15 scales (37%), for example, the Multi-ligament Quality of Life questionnaire [71] and Quality of Life Questionnaire Outcome Measure for Chronic Anterior Cruciate Ligament (ACL) Deficiency [9]. Among these, the level of frustration, fear, and sense of security, etc. is assessed. In fact, one for ACL injuries is based exclusively on psycho-emotional data [10]. It seems contradictory to incorporate subjective aspects into tools that seek to quantify and objectify the functional assessment of users but

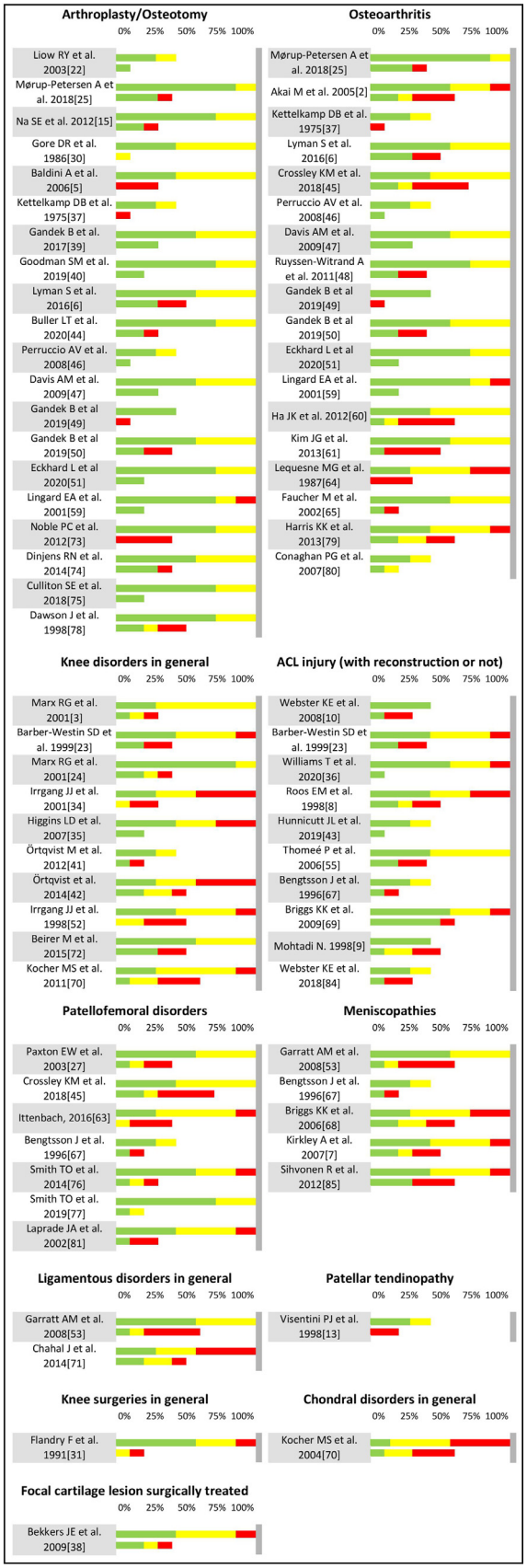

Fig. 3. Comparison of methodological quality of the validation studies according to population groups. * QUADAS-2 y COSMIN-RB are represented in the first and second bar, respectively. Green = "low risk of bias" or "low concerns regarding applicability" (QUADAS-2); "very good" and "adequate" (COSMIN RB). Yellow = "unclear risk of bias" or "unclear concerns regarding applicability" (QUADAS-2); "doubtful" (COSMIN RB). Red = "high risk of bias" or "high concerns regarding applicability" (QUADAS-2); "inadequate" (COSMIN RB). White = items that could not be evaluated (QUADAS-2); metric properties not evaluated (COSMIN RB). ACL, anterior cruciate ligament; QUADAS-2, Quality Assessment of Diagnostic Accuracy Studies; COSMIN-RB, COnsensus-based Standards for the selection of health Measurement Instruments Risk of Bias.

is completely necessary because psychological and physical aspects must be treated together for holistic and effective assessments [107].

4.4. Age of users

Only 2 scales were specific for patients < 18 years old, but for general disorders: the KOOS-Child [41] and PEDI-IKDC [82]. This situation is obvious because of the high knee pathology prevalence in adulthood. Thus, joint degeneration, and with it, osteoarthritis [108] and total knee arthroplasties [109], especially those that are consolidated [110], are frequent in older people. Also, accidental injuries are more associated with the work environment [111], including professional sports, and with traffic accidents, although they are not exclusive to adults. Accordingly, most of the validated scales were used in osteoarthritis, arthroplasty and ACL injury.

Thus, there is less need for assessment in childhood. However, pathologies before adulthood should be considered, such as Osgood-Schlatter and Sinding Larsen Johansson disease in adolescents [112], congenital misalignments (i.e., genu valgum or varum [113]), bipartite patella [112], and recurrent patellar dislocations [114]. Hence, more validated tools for particular knee disorders in children and with higher methodological quality than those published to date are recommended.

4.5. Tool administration

Many tools used informal and simple language to facilitate user understanding. Some questions such as "Do you have problems going up stairs?" [31] or "How often does your knee make you limp?" [71] can be found. Even 3 scales, the KOOS-Child [41], MKQ [72] and VISA-P [13], use images representing actions, gestures or positions, an essential aspect when addressing children [41]. Regardless, many tools depend on clinicians specialized in knees to apply them because they assess clinical signs; physical examination tests such as the Lachman test [33]; the varus/valgus, anterior/posterior drawer; or diagnostic tests as radiography [4,33]. For example, the International Knee Documentation Committee [33] tool contains image results with millimetric variations; the Cincinnati Knee Rating [4] rates spillage based on cubic centimeters; and the KSS [58] tests ligament stability. From previous reasoning and authors' suggestions, 30 of 41 tools were self-administered. Those have the advantage of considering the patient's perceptions (patient-reported outcome measure), and improving user–clinician communication. In addition, self-administered tools do not require the presence of users or clinicians. However, they do not consider the clinician's perception and cannot include physical tests with objective data, in contrast to non–self-administered scales.

We consider that there are no better or worse assessment tools. The clinical circumstances, evaluation context and patient characteristics could lead clinicians to select a lower methodological quality tool, in favor of other aspects such as population specificity, contents covered or type of administration [115].

5. Limitations

The limitations of this study were the use of reference managers and the screening of records owing to the large number of records (> 340,000). Also, the third column in Appendix A has a limitation in

that it includes a maximum of only 6 indications/applications. Including all of them would mean exponentially increasing both the extension and references of the review. This is the starting point for prospective systematic reviews addressing all implementations, transcultural adaptations of each scale, and even reviews of tool subgroups.

6. Strengths

This review brought together validated outcome measures to assess knees that are unlimited in time. When assessments by outcome measures are needed, some of the difficulties for clinicians are not knowing which ones exist, how to access them, choosing the most appropriate in each case, not knowing the application instructions and interpreting the score. This study provides descriptive data for the content of each scale, possible applications and useful and necessary information for using the tool. Also, because the physical scales cannot be shown so as not to incur plagiarism, guidelines and references are marked for easily locating them. The review facilitates and encourages the initial, continuous and final assessments of the functional recovery processes. Consequently, it enables the evaluation of the user's progression, comparisons between interventions and the creation of effective clinical action protocols by improving clinical practice. We also add the strength of an exhaustive analysis of the methodological quality of the tool validations, guaranteeing clinicians and/or researchers that tools/scales were created based on scientific evidence.

Therefore, this review is a particularly useful and feasible tool in daily clinical practice to assess the functional knee recovery processes of users.

7. **Conclusions** and implications

We conducted a necessary compilation of validated outcome measures of functional knee recovery that was unlimited by time. We found limited methodological quality in the 71 validation studies (original and complementary) related to the 41 tools analysed. The best validations were for the CKRS in osteoarthritis and arthroplasties, the KOS-ADL and Lysholm Knee Score for knee disorders in general, and the TAS for ACL injuries. This operational document provides data on outcome measures, parameter descriptions, implementations, instructions, interpretations and languages of the scales, thus allowing to identify the most appropriate tool to be used in future interventions. The most approached topics were specific knee functions, especially gait, pain/sensitivity and physical activity and sports.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author's contributions

GC conceptualized the idea and supervised the teamwork. GC, FE CR and DT performed data curation, investigation, formal analysis and writing - original draft. GC, FE, VP and CR critically carried out writing - review & editing. All authors approved the final paper for publication.

Declaration of Competing interest

None declared.

Acknowledgements

The authors thank the Research Group "Area of Physiotherapy CTS-305″ of the University of Seville, Spain for its contribution in this study.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.rehab.2021.101608.

References

[1] Hol AM, van Grinsven S, Lucas C, van Susante JLC, van Loon CJM. Partial versus unrestricted weight bearing after an uncemented femoral stem in total hip arthroplasty: recommendation of a concise rehabilitation protocol from a systematic review of the literature. Arch Orthop Trauma Surg 2010;130:547–55.
[2] Akai M, Doi T, Fujino K, Iwaya T, Kurosawa H, Nasu T. An outcome measure for Japanese people with knee osteoarthritis. J Rheumatol 2005;32:1524–32.
[3] Marx RG, Stump TJ, Jones EC, Wickiewicz TL, Warren RF. Development and evaluation of an activity rating scale for disorders of the knee. Am J Sports Med 2001;29:213–8.
[4] Noyes FR, Mooar PA, Matthews DS, Butler DL. The symptomatic anterior cruciate-deficient knee. Part I: the long-term functional disability in athletically active individuals. J Bone Jt Surg - Ser A 1983;65:154–62.
[5] Baldini A, Anderson JA, Zampetti P, Pavlov H, Sculco TP. A new patellofemoral scoring system for total knee arthroplasty. Clin Orthop Relat Res 2006:150–4.
[6] Lyman S, Lee Y-Y, Franklin PD, Li W, Cross MB, Padgett DE. Validation of the KOOS, JR: a short-form knee arthroplasty outcomes survey. Clin Orthop Relat Res 2016;474:1461–71.
[7] Kirkley A, Griffin S, Whelan D. The development and validation of a quality of life-measurement tool for patients with meniscal pathology: the Western Ontario Meniscal Evaluation Tool (WOMET). Clin J Sport Med 2007;17:349–56.
[8] Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee injury and osteoarthritis outcome score (KOOS) - development of a self-administered outcome measure. J Orthop Sports Phys Ther 1998;28:88–96.
[9] Mohtadi N. Development and validation of the quality of life outcome measure (questionnaire) for chronic anterior cruciate ligament deficiency. Am J Sports Med 1998;26:350–7.
[10] Webster KE, Feller JA, Lambros C. 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 2008;9:9–15.
[11] Williams G, Morris ME, Schache A, McCrory P. Observational gait analysis in traumatic brain injury: accuracy of clinical judgment. Gait Posture 2009;29:454–9.
[12] Chamorro-Moriana G, Ridao-Fernández C, Ojeda J, Benítez-Lugo M, Sevillano JL. Reliability and validity study of the chamorro assisted gait scale for people with sprained ankles, walking with forearm crutches. PLoS ONE 2016;11:e0155225.
[13] Visentini PJ, Khan KM, Cook JL, Kiss ZS, Harcourt PR, Wark JD. The VISA score: an index of severity of symptoms in patients with jumper's knee (Patellar Tendinosis). J Sci Med Sport 1998;1:22–8.
[14] Reid S, Held JM, Lawrence S. Reliability and validity of the Shaw gait assessment tool for temporospatial gait assessment in people with hemiparesis. Arch Phys Med Rehabil 2011;92:1060–5.
[15] Na SE, Ha CW, Lee CH. A new high-flexion knee scoring system to eliminate the ceiling effect. Clin Orthop Relat Res 2012;470:584–93.
[16] Gagnier JJ, Mullins M, Huang H, Marinac-Dabic D, Ghambaryan A, Eloff B, et al. A systematic review of measurement properties of patient-reported outcome measures used in patients undergoing total knee arthroplasty. J Arthroplasty 2017;32:1688–97 e7.
[17] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 2009;62:e1–34.
[18] Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. Quadas-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529–36.
[19] Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. Qual Life Res 2018;27:1171–9.
[20] Mokkink L, B; Terwee CB;, Patrick DL;, Alonso J;, Stratford PW;, Knol DL;, Bouter L M; De, Vet HCW. COSMIN checklist manual. 2012 http://cosmin.nl.
[21] Aichroth P, Freeman MAR, Smillie IS, Souter WA. A knee function assessment chart. J Bone Jt Surg - Ser B 1978;60 B:308–9.
[22] Liow RYL, Walker K, Wajid MA, Bedi G, Lennox CME. Functional rating for knee arthroplasty: comparison of three scoring systems. Orthopedics 2003;26:143–9.
[23] Barber-Westin SD, Noyes FR, McCloskey JW. 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. Am J Sports Med 1999;27:402–16.
[24] Marx RG, Jones EC, Allen AA, Altchek DW, O'Brien SJ, Rodeo SA, et al. Reliability, validity, and responsiveness of four knee outcome scales for athletic patients. J Bone Jt Surg - Ser A 2001;83:1459–69.

[25] Mørup-Petersen A, Holm PM, Holm CE, Klausen TW, Skou ST, Krogsgaard MR, et al. Knee osteoarthritis patients can provide useful estimates of passive knee range of motion: development and validation of the copenhagen knee ROM scale. J Arthroplasty 2018;33 2875–83.e3.
[26] Fulkerson JP, Becker GJ, Meaney JA, Miranda M, Folcik MA. Anteromedial tibial tubercle transfer without bone graft. Am J Sports Med 1990;18:490–7.
[27] Paxton EW, Fithian DC, Stone ML, Silva P. The reliability and validity of knee-specific and general health instruments in assessing acute patellar dislocation outcomes. Am J Sports Med 2003;31:487–92.
[28] Cerciello S, Corona K, Morris BJ, Visonà E, Maccauro G, Maffulli N, et al. Cross-cultural adaptation and validation of the Italian versions of the Kujala, Larsen, Lysholm and Fulkerson scores in patients with patellofemoral disorders. J Orthop Traumatol 2018;19:18.
[29] Ranawat CS, Shine JJ. Duo-condylar total knee arthroplasty. Clin Orthop Relat Res 1973:185–95.
[30] Gore DR, Murray MP, Sepic SB, Gardner GM. Correlations between objective measures of function and a clinical knee rating scale following total knee replacement. Orthopedics 1986;9:1363–7.
[31] Flandry F, Hunt JP, Terry GC, Hughston JC. Analysis of subjective knee complaints using visual analog scales. Am J Sports Med 1991;19:112–8.
[32] Hooper DM, Morrissey MC, Drechsler WI, McDermott M, McAuliffe TB. Validation of the hughston clinic subjective knee questionnaire using gait analysis. Med Sci Sports Exerc 2001;33:1456–62.
[33] Hefti E, Müller W, Jakob RP, Stäubli HU. Evaluation of knee ligament injuries with the IKDC form. Knee Surg Sport Traumatol Arthrosc 1993;1:226–34.
[34] Irrgang JJ, Anderson AF, Boland AL, Harner CD, Kurosaka M, Neyret P, et al. Development and validation of the International Knee Documentation Committee Subjective Knee form. Am J Sports Med 2001;29:600–13.
[35] Higgins LD, Taylor MK, Park D, Ghodadra N, Marchant M, Pietrobon R, et al. Reliability and validity of the International Knee Documentation Committee (IKDC) Subjective Knee Form. Jt Bone Spine 2007;74:594–9.
[36] Williams T, Burley D, Evans L, Robertson A, Hardy L, Roy S, et al. The structural validity of the IKDC and its relationship with quality of life following ACL reconstruction. Scand J Med Sci Sport 2020;30:1748–57.
[37] Kettelkamp DB, Thompson C. Development of a knee scoring scale. Clin Orthop 1975;107:93–9.
[38] Bekkers JEJ, de Windt TS, Raijmakers NJH, Dhert WJA, Saris DBF. Validation of the Knee Injury and Osteoarthritis Outcome Score (KOOS) for the treatment of focal cartilage lesions. Osteoarthr Cartil 2009;17:1434–9.
[39] Gandek B, Ware Jr. JE. Validity and responsiveness of the knee injury and osteoarthritisoutcome score: a comparative study among total knee replacement patients. Arthritis Care Res (Hoboken) 2017;69:817–25.
[40] Goodman SM, Mehta BY, Mandl LA, Szymonifka JD, Finik J, Figgie MP, et al. Validation of the hip disability and osteoarthritis outcome score and knee injury and osteoarthritis outcome score pain and function subscales for use in total hip replacement and total knee replacement clinical trials. J Arthroplasty 2020;35 1200–7.e4.
[41] Örtqvist M, Roos EM, Broström EW, Janarv P-M, Iversen MD. Development of the Knee Injury and Osteoarthritis Outcome Score for Children (KOOS-Child): comprehensibility and content validity. Acta Orthop 2012;83:666–73.
[42] Örtqvist M, Iversen MD, Janarv P-M, Broström EW, Roos EM. Psychometric properties of the knee injury and osteoarthritis outcome score for children (KOOS-Child) in children with knee disorders. Br J Sports Med 2014;48: 1437–46.
[43] Hunnicutt JL, Hand BN, Gregory CM, Slone HS, McLeod MM, Pietrosimone B, et al. KOOS-JR demonstrates psychometric limitations in measuring knee health in individuals after ACL reconstruction. Sports Health 2019;11:242–6.
[44] Buller LT, McLawhorn AS, Lee YY, Cross M, Haas S, Lyman S. The short form KOOS, JR is valid for revision knee arthroplasty. J Arthroplasty 2020;35:2543–9.
[45] Crossley KM, Macri EM, Cowan SM, Collins NJ, Roos EM. The patellofemoral pain and osteoarthritis subscale of the KOOS (KOOS-PF): development and validation using the COSMIN checklist. Br J Sports Med 2018;52:1130–6.
[46] Perruccio AV, Stefan Lohmander L, Canizares M, Tennant A, Hawker GA, Conaghan PG, et al. The development of a short measure of physical function for knee OA KOOS-Physical Function Shortform (KOOS-PS) - an OARSI/OMERACT initiative. Osteoarthr Cartil 2008;16:542–50.
[47] Davis AM, Perruccio AV, Canizares M, Hawker GA, Roos EM, Maillefert JF, et al. Comparative, validity and responsiveness of the HOOS-PS and KOOS-PS to the WOMAC physical function subscale in total joint replacement for Osteoarthritis. Osteoarthr Cartil 2009;17:843–7.
[48] Ruyssen-Witrand A, Fernandez-Lopez CJ, Gossec L, Anract P, Courpied JP, Dougados M. Psychometric properties of the OARSI/OMERACT osteoarthritis pain and functional impairment scales: ICOAP, KOOS-PS and HOOS-PS. Clin Exp Rheumatol 2011;29:231–7.
[49] Gandek B, Roos EM, Franklin PD, Ware JE. Item selection for 12-item short forms of the knee injury and osteoarthritis outcome score (KOOS-12) and hip disability and osteoarthritis outcome score (HOOS-12). Osteoarthr Cartil 2019;27:746–53.
[50] Gandek B, Roos EM, Franklin PD, Ware Jr. JE. A 12-item short form of the knee injury and osteoarthritis outcome score(KOOS-12): tests of reliability, validity and responsiveness. Osteoarthr Cartil 2019;27:762–70.
[51] Eckhard L, Munir S, Wood D, Talbot S, Brighton R, Walter B, et al. The KOOS-12 shortform shows no ceiling effect, good responsiveness and construct validity compared to standard outcome measures after total knee arthroplasty. Knee Surg Sport Traumatol Arthrosc 2021;29:608–15.
[52] Irrgang JJ, Snyder-Mackler L, Wainner RS, Fu FH, Harner CD. Development of a patient-reported measure of function of the knee. J Bone Joint Surg Am 1998;80:1132–45.
[53] Garratt AM, Brealey S, Robling M, Atwell C, Russell I, Gillespie W, et al. Development of the knee quality of life (KQoL-26) 26-item questionnaire: data quality, reliability, validity and responsiveness. Health Qual Life Outcomes 2008;6:48.
[54] Chuang LH, Garratt A, Brealey S. Comparative responsiveness and minimal change of the knee quality of life 26-item (KQoL-26) questionnaire. Qual Life Res 2013;22:2461–75.
[55] Thomeé P, Währborg P, Börjesson M, Thomeé R, Eriksson BI, Karlsson J. A new instrument for measuring self-efficacy in patients with an anterior cruciate ligament injury. Scand J Med Sci Sport 2006;16:181–7.
[56] Ericsson YB, Ringsberg K, Dahlberg LE. Self-efficacy, physical activity and health-related quality of life in middle-aged meniscectomy patients and controls. Scand J Med Sci Sport 2011;21:e150–8.
[57] van Lankveld W, van Melick N, Habets B, Pronk Y, Bart Staal J, van Cingel R. Cross-cultural adaptation and measurement properties of the Dutch knee self efficacy scale (K-SES). BMC Sports Sci Med Rehabil 2019;11:3.
[58] Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. Clin Orthop Relat Res 1989:13–4.
[59] Lingard EA, Katz JN, Wright RJ, Wright EA, Sledge CB. Validity and responsiveness of the knee society clinical rating system in comparison with the SF-36 and WOMAC. J Bone Jt Surg - Ser A 2001;83:1856–64.
[60] Ha JK, Kim JG, Lee MC, Wang JH. What symptoms are more important for korean patients in knee osteoarthritis? Development and validation of the Korean Knee score. Knee Surg Relat Res 2012;24:151–7.
[61] Kim JG, Ha JK, Han SB, Kim TK, Lee MC. Development and validation of a new evaluation system for patients with a floor-based lifestyle: the Korean knee score knee. Clin Orthop Relat Res 2013;471:1539–47.
[62] Kujala UM, Jaakkola LH, Koskinen SK, Taimela S, Hurme M, Nelimarkka O. Scoring of patellofemoral disorders. Arthroscopy 1993;9:159–63.
[63] Ittenbach RF, Huang G, Foss KDB, Hewett TE, Myer GD. Reliability and validity of the anterior knee pain scale: applications for use as an epidemiologic screener. PLoS ONE 2016;11:e0159204.
[64] Lequesne MG, Mery C, Samson M, Gerard P. Indexes of severity for osteoarthritis of the hip and knee: validation-value in comparison with other assessment tests. Scand J Rheumatol 1987;16:85–9.
[65] Faucher M, Poiraudeau S, Lefevre-Colau MM, Rannou F, Fermanian J, Revel M. Algo-functional assessment of knee osteoarthritis: comparison of the test-retest reliability and construct validity of the Womac and Lequesne indexes. Osteoarthr Cartil 2002;10:602–10.
[66] Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. Clin Orthop Relat Res 1985;198:43–9.
[67] Bengtsson J, Möllborg J, Werner S. A study for testing the sensitivity and reliability of the Lysholm knee scoring scale. Knee Surgery, Sport Traumatol Arthrosc 1996;4:27–31.
[68] Briggs KK, Kocher MS, Rodkey WG, Steadman JR. Reliability, validity, and responsiveness of the Lysholm knee score and Tegner activity scale for patients with meniscal injury of the knee. J Bone Joint Surg Am 2006;88:698–705.
[69] Briggs KK, Lysholm J, Tegner Y, Rodkey WG, Kocher MS, Steadman JR. The reliability, validity, and responsiveness of the Lysholm score and Tegner activity scale for anterior cruciate ligament injuries of the knee: 25 years later. Am J Sports Med 2009;37:890–7.
[70] Kocher MS, Steadman JR, Briggs KK, Sterett WI, Hawkins RJ. Reliability, validity, and responsiveness of the Lysholm knee scale for various chondral disorders of the knee. J Bone Jt Surg - Ser A 2004;86:1139–45.
[71] Chahal J, Whelan DB, Jaglal SB, Smith P, Macdonald PB, Levy BA, et al. The multiligament quality of life questionnaire: development and evaluation of test-retest reliability and validity in patients with multiligament knee injuries. Am J Sports Med 2014;42:2906–16.
[72] Beirer M, Fiedler N, Huber S, Schmitt-Sody M, Lorenz S, Biberthaler P, et al. The munich knee questionnaire: development and validation of a new patient-reported outcome measurement tool for knee disorders. Arthrosc - J Arthrosc Relat Surg 2015;31:1522–9.
[73] Noble PC, Scuderi GR, Brekke AC, Sikorskii A, Benjamin JB, Lonner JH, et al. Development of a new Knee Society scoring system. Clin Orthop Relat Res 2012;470:20–32.
[74] Dinjens RN, Senden R, Heyligers IC, Grimm B. Clinimetric quality of the new 2011 knee society score: high validity, low completion rate. Knee 2014;21:647–54.
[75] Culliton SE, Bryant DM, MacDonald SJ, Hibbert KM, Chesworth BM. Validity and internal consistency of the new knee society knee scoring system. Clin Orthop Relat Res 2018;476:77–84.
[76] Smith TO, Donell ST, Clark A, Chester R, Cross J, Kader DF, et al. The development, validation and internal consistency of the Norwich Patellar Instability (NPI) score. Knee Surg Sport Traumatol Arthrosc 2014;22:324–35.
[77] Smith TO, Choudhury A, Navratil R, Hing CB. Psychometric properties of the Norwich Patellar Instability Score in people with recurrent patellar dislocation. Knee 2019;26:1192–7.
[78] Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. J Bone Jt Surg - Ser B 1998;80:63–9.
[79] Harris KK, Dawson J, Jones LD, Beard DJ, Price AJ. Extending the use of PROMs in the NHS-using the Oxford Knee Score in patients undergoing non-operative management for knee osteoarthritis: a validation study. BMJ Open 2013;3:e003365.
[80] Conaghan PG, Emerton M, Tennant A. Internal construct validity of the Oxford knee scale: evidence from Rasch measurement. Arthritis Care Res 2007;57:1363–7.