Osteoarthritis and Cartilage xxx (2016) 1–13

Osteoarthritis and Cartilage

OARSI OSTEOARTHRITIS RESEARCH SOCIETY INTERNATIONAL

Review

Knee Injury and Osteoarthritis Outcome Score (KOOS): systematic review and meta-analysis of measurement properties

N.J. Collins † ‡ *, C.A.C. Prinsen §, R. Christensen ||, E.M. Bartels ¶, C.B. Terwee §, E.M. Roos #

† Division of Physiotherapy, School of Health and Rehabilitation Sciences, The University of Queensland, St Lucia, Queensland, Australia
‡ Department of Mechanical Engineering, Melbourne School of Engineering, The University of Melbourne, Parkville, Victoria, Australia
§ VU University Medical Center, Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, The Netherlands
|| Musculoskeletal Statistics Unit, The Parker Institute, Department of Rheumatology, Copenhagen University Hospital, Bispebjerg and Frederiksberg Hospital, Frederiksberg, Denmark
¶ The Biochemistry and Physiology Laboratory, The Parker Institute, Department of Rheumatology, Copenhagen University Hospital, Bispebjerg and Frederiksberg, Frederiksberg, Denmark
# Research Unit for Musculoskeletal Function and Physiotherapy, Institute of Sports Science and Clinical Biomechanics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark

ARTICLE INFO

Article history:
Received 24 August 2015
Accepted 13 March 2016

Keywords:
Knee injury
Knee osteoarthritis
KOOS
Patient-reported outcome measures
Measurement properties
Meta-analysis

SUMMARY

Objective: To conduct a systematic review and meta-analysis to synthesize evidence regarding measurement properties of the Knee Injury and Osteoarthritis Outcome Score (KOOS).
Design: A comprehensive literature search identified 37 eligible papers evaluating KOOS measurement properties in participants with knee injuries and/or osteoarthritis (OA). Methodological quality was evaluated using the COSMIN checklist. Where possible, meta-analysis of extracted data was conducted for all studies and stratified by age and knee condition; otherwise narrative synthesis was performed.
Results: KOOS has adequate internal consistency, test-retest reliability and construct validity in young and old adults with knee injuries and/or OA. The ADL subscale has better content validity for older patients and Sport/Rec for younger patients with knee injuries, while the Pain subscale is more relevant for painful knee conditions. The five-factor structure of the original KOOS is unclear. There is some evidence that the KOOS subscales demonstrate sufficient unidimensionality, but this requires confirmation. Although measurement error requires further evaluation, the minimal detectable change for KOOS subscales ranges from 14.3 to 19.6 for younger individuals, and ≥20 for older individuals. Evidence of responsiveness comes from larger effect sizes following surgical (especially total knee replacement) than non-surgical interventions.
Conclusions: KOOS demonstrates adequate content validity, internal consistency, test-retest reliability, construct validity and responsiveness for age- and condition-relevant subscales. Structural validity, cross-cultural validity and measurement error require further evaluation, as well as construct validity of KOOS-PS. Suggested order of subscales for different knee conditions can be applied in hierarchical testing of endpoints in clinical trials.
Systematic review registration: PROSPERO (CRD42011001603).

Introduction

Patient-reported outcome measures (PROMs) are used across medical disciplines to follow disease course or evaluate treatment outcomes. PROMs involve the patient's evaluation of any aspect of their health status, without interpretation of their response by another individual[1]. Study findings are more truthful if the PROM used has adequate measurement properties, and if the studies testing its measurement properties are of good or excellent

* Address correspondence and reprint requests to: N. Collins, School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane QLD 4072, Australia.
E-mail addresses: n.collins1@uq.edu.au (N.J. Collins), c.prinsen@vumc.nl (C.A.C. Prinsen), robin.christensen@regionh.dk (R. Christensen), Else.Marie.Bartels@regionh.dk (E.M. Bartels), cb.terwee@vumc.nl (C.B. Terwee), eroos@health.sdu.dk (E.M. Roos).

http://dx.doi.org/10.1016/j.joca.2016.03.010
1063-4584/

methodological quality[2]. Specifically, a PROM should have content that is relevant for the construct of interest and the target population, measure intended dimensions, be stable on repeated measures, and be able to detect change in patients' perceived health status[3]. Since measurement properties can differ between different patient groups (e.g., patient characteristics, medical condition, intervention), PROMs need to be evaluated in different patient populations, ideally across multiple studies.

The Knee Injury and Osteoarthritis Outcome Score (KOOS) is a PROM intended for young, middle-aged and elderly adults with knee injury and/or knee osteoarthritis (OA), and can be used to monitor disease course and outcomes following surgical, pharmacological and other interventions[4]. KOOS holds five subscales: (1) Pain (9 items); (2) other Symptoms (7 items); (3) Activities of Daily Living (ADL, 17 items); (4) Sport and Recreation function (Sport/Rec, 5 items); and (5) knee-related Quality of Life (QoL, 4 items). Each subscale is scored separately from zero (extreme knee problems) to 100 (no knee problems). The KOOS Physical function Short form (KOOS-PS, 7 items) was later derived from the ADL and Sport/Rec subscales via Rasch analysis[5]. The clinical and research utility of KOOS is highlighted by large international patient datasets (>100,000 unique patient records) and frequent use in scientific publications. Importantly, KOOS has international accessibility, being free of charge and translated into >45 different language versions[6].

To provide clinicians and researchers with a single reference regarding KOOS measurement properties, we performed a systematic review and meta-analysis to evaluate the measurement properties of KOOS in people with knee injuries and/or OA.

Method

Review protocol

The protocol was developed according to the PRISMA statement[7] and prospectively registered (PROSPERO, CRD42011001603, 11 October 2011).

Literature search

A research librarian and researcher in arthritic diseases (E.M.B.) searched six bibliographic databases from 1998 (date of KOOS publication) to 16 January 2014, with no language restrictions (Medline via PubMed, EMBASE via OVID, CINAHL via EBSCO, Web of Science, Psycinfo via OVID, Cochrane Central Register of Controlled Trials). The following terms were searched as free text and key words (where applicable): (KOOS AND knee) OR (Knee Injury and Osteoarthritis Outcome Score). Reference lists of eligible papers and review papers were manually searched. The KOOS website[6] was reviewed to cross-check inclusion of all cross-cultural validation studies or papers published in non-English languages, and KOOS developers were consulted regarding known unpublished studies.

Eligibility criteria

Original full-text published studies were considered for inclusion, as well as PhD theses identified by the search strategy. Studies were eligible if: (1) the primary aim was to evaluate at least one measurement property (e.g., reliability, validity, responsiveness) or interpretability of KOOS/KOOS-PS; (2) they studied participants of any age suffering from any knee injury and/or knee OA; and (3) KOOS was patient-completed (e.g., paper, computer or touch-screen administration). Where appropriate, we included studies that utilized a research administrator to assist KOOS completion in populations with limited education (e.g., interviewer-administered), but excluded studies where medical practitioners administered KOOS in the clinic to reduce the risk of bias. No restrictions were placed on method of study recruitment, study venue, or KOOS language version. Studies were excluded if they used KOOS to assess participants for which KOOS was not designed (e.g., other lower limb conditions, asymptomatic cohorts), to evaluate treatment efficacy without assessing measurement properties, or to validate other measurement tools.

Article selection

Results of database searches were imported into Reference Manager 12 (Thomson Reuters, Philadelphia, USA). Two independent reviewers (N.J.C., E.M.R.) assessed titles, abstracts and full-text articles (where appropriate) for eligibility. Discrepancies were discussed to reach consensus, with unresolved cases taken to a third reviewer (C.B.T.).

Data extraction

One reviewer (N.J.C.) used a predefined spreadsheet to extract study characteristics and measurement property data on two separate occasions, blinded to the previous extraction[8]. There were minimal discrepancies between the two extractions, resolved by a second reviewer (R.C.). For studies published in languages other than English, we contacted corresponding authors to assist with translation of required data. When authors were unavailable, we utilized university volunteers who were native speakers.

Evaluation of measurement properties

We evaluated specific measurement properties as defined by the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) taxonomy[9,10]: content validity, structural validity and unidimensionality, internal consistency, test-retest reliability, measurement error, construct validity, cross-cultural validity and responsiveness (Table I). We also evaluated interpretability and feasibility. Criterion validity was not evaluated due to the lack of an established gold standard measure for domains captured by KOOS[13].

Data synthesis

For test-retest reliability, we calculated weighted mean intra-class correlation coefficients (ICCs) and 95% confidence intervals using a standard generic inverse variance random effects model[16]. ICC values were combined based on estimates derived from a Fisher transformation, $z = 0.5 \times \ln((1 + \text{ICC})/(1-\text{ICC}))$, which has an approximate variance, $(\text{Var}(z) = 1/(N\text{-}3))$, where N is the sample size. For each analysis, we qualitatively evaluated whether it was appropriate to pool data (eyeballing), and used quantitative methods (I^2) to evaluate the potential impact of heterogeneity on pooled estimates[17,18]. The likelihood of publication bias was addressed qualitatively, by contacting knee OA experts with an interest in psychometrics. We performed subgroup analyses for different age groups and knee conditions, comparing strata using meta-regression with study characteristics handled as independent class variables (SAS software, version 9.3).

For other measurement properties, we calculated weighted means (number of participants included per study) and 95% confidence intervals, where possible and appropriate (qualitative evaluation), as well as subgroup analyses for different age groups and knee conditions. Quantitative findings for each measurement

Please cite this article in press as: Collins NJ, et al., Knee Injury and Osteoarthritis Outcome Score (KOOS): systematic review and meta-analysis of measurement properties, Osteoarthritis and Cartilage (2016), http://dx.doi.org/10.1016/j.joca.2016.03.010