



■ INSTRUCTIONAL REVIEW: KNEE

Patient-reported outcome measures after total knee arthroplasty

A SYSTEMATIC REVIEW

P. N. Ramkumar,
J. D. Harris,
P. C. Noble

From Houston
Methodist
Orthopedics and
Sports Medicine,
Houston, Texas,
United States

Objectives

A lack of connection between surgeons and patients in evaluating the outcome of total knee arthroplasty (TKA) has led to the search for the ideal patient-reported outcome measure (PROM) to evaluate these procedures. We hypothesised that the desired psychometric properties of the ideal outcome tool have not been uniformly addressed in studies describing TKA PROMS.

Methods

A systematic review was conducted investigating one or more facets of patient-reported scores for measuring primary TKA outcome. Studies were analysed by study design, subject demographics, surgical technique, and follow-up adequacy, with the 'gold standard' of psychometric properties being systematic development, validity, reliability, and responsiveness.

Results

A total of 38 articles reported outcomes from 47 different PROMS to 85 541 subjects at 26.3 months (standard deviation 30.8) post-operatively. Of the 38, eight developed new scores, 20 evaluated existing scores, and ten were cross-cultural adaptation of existing scores. Only six of 38 surveyed studies acknowledged all 'gold standard' psychometric properties. The most commonly studied PROMS were the Oxford Knee Score, New Knee Society Score, Osteoarthritis Outcome Score, and Western Ontario and McMaster Universities Osteoarthritis Index.

Conclusions

A single, validated, reliable, and responsive PROM addressing TKA patients' priorities has not yet been identified. Moreover, a clear definition of a successful procedure remains elusive.

Cite this article: *Bone Joint Res* 2015;4:120–127

■ P. N. Ramkumar, MBA, Medical Student
Hospital for Special Surgery, 535
E 70th St, Bellaire Building, Suite
9-11, New York 10021, USA.

■ J. D. Harris, MD, Orthopaedic Surgeon
■ P. C. Noble, PhD, Professor
Houston Methodist Orthopedics and Sports Medicine,
6550 Fannin Street,
Smith Tower, Suite 2511
Houston, Texas 77030, USA.

Correspondence should be sent to Mr P. N. Ramkumar; email: premrnkumar@gmail.com

doi:10.1302/2046-3758.47.2000380 \$2.00

2015;4: 120 – 127.
Received 08 January 2015;
Accepted after revision 19 March 2015

Keywords: Outcomes, TKA, Measure

Introduction

The value of total knee arthroplasty (TKA) is being subjected to increased scrutiny, due both to the cost of the procedure to the healthcare system, and to numerous reports suggesting disturbing rates of patient dissatisfaction with the outcome of this procedure.¹⁻⁸ Moreover, given the ubiquity of osteoarthritis of the knee and the demand for prosthetic replacement, increasing pressure is being placed on healthcare providers to demonstrate the benefits and cost effectiveness of TKA in an objective manner.⁹⁻¹¹ This has prompted a recent push to quantify the

success of these procedures using patient-reported outcome measures (PROMS).¹²⁻¹⁴

Like many medical procedures, any evaluation of value and effectiveness of TKA depends on our definition of a 'successful' treatment of a debilitating condition within the multiple dimensions of clinical outcome. These include relief of the patient's symptoms, restoration of their physical function, and realisation of their pre- and post-operative aspirations.²⁻⁶ Previous studies have shown that 10% to 25% of patients are dissatisfied with the outcome of knee replacement at one to three years

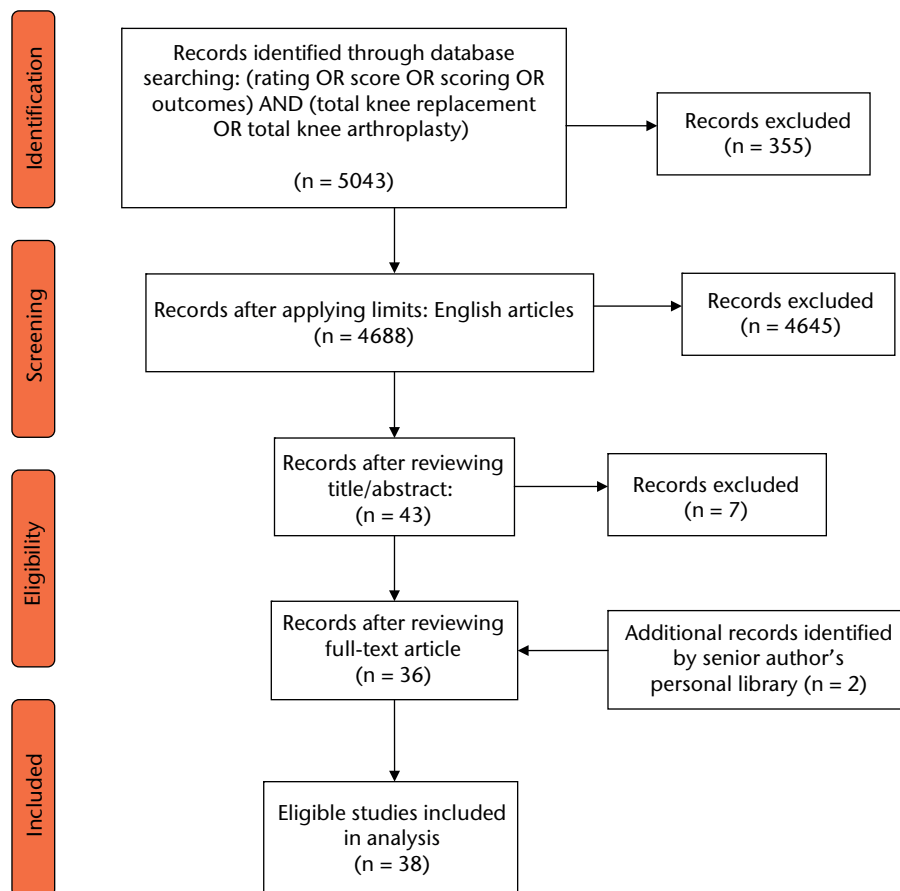


Fig. 1

PRISMA flow diagram.

after surgery, primarily because of the presence of residual symptoms and the failure of the procedure to meet the patient's pre-operative expectations.^{6-8,11,13-15}

An additional disturbing finding is the difference in perception of patients and their surgeons regarding the success of the procedure, with the view of surgeons often being more sanguine.^{2,7,8,11,15,21} However, these expectations are both multifaceted and evolve as part of the process of recovery and psychological adjustment following surgery. The immediate expectation of patients undergoing joint arthroplasty is relief of pain.¹⁶⁻²⁰ Moreover, almost all patients expect to be able to perform activities of daily living (ADLs) without limitation after surgery. Unfortunately, for nearly all clinical conditions, including TKA, a single, properly developed, validated, reliable, and responsive PROM has not yet been developed.^{15,21}

A systematic review was conducted to elucidate the psychometric properties and variations among the available PROMS. The review sought to answer three questions:

- Which PROMs designed to measure the outcome of TKA have been shown to be valid, reliable, responsive and systematically developed?

- What do these instruments measure?

- Is one instrument best suited to measuring the functional status of patients with knee complaints and at follow-up after TKA?

We hypothesised that the desired psychometric properties (development, validity, reliability, and responsiveness) of the ideal outcome instrument have not been uniformly addressed in all the studies analysing PRO instruments.

Materials and Methods

A systematic review was conducted in accordance with Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines to address the hypotheses of this study. Prior to the commencement, the systematic review protocol was registered using the PROSPERO database (CRD42013005604; September 4, 2013). Two independent reviewers separately completed the search. The initial search was performed on October 29, 2014 using the Medline database, with cross-checking of reference lists for inclusion. Search terms were: "(rating OR score OR scoring OR outcomes) AND (total knee replacement OR total knee arthroplasty OR TKA OR

TKR)". All studies with levels of evidence ranging from I to IV (according to the Oxford Centre for Evidence-Based Medicine) were included.²² Publications in the past ten years were largely selected, however, commonly referenced and highly regarded older publications were not excluded. Articles meeting these criteria were manually reviewed and were discussed among the authors, and a decision was made regarding inclusion. If there was any disagreement among authors regarding the inclusion of an article, the final decision was made by consensus or, if necessary, the decision of the senior author. A PRISMA flowchart was created for transparency of article selection (Fig. 1). The heterogeneity of studies and their associated outcomes precluded the formulation of a meta-analysis.

Inclusion criteria. A study describing an outcome score that possessed proper development, validity, reliability, and responsiveness was defined as the 'gold standard'. To meet inclusion in this review, articles were required to investigate one or more of the following facets of patient-reported scores for measuring the outcome of primary total knee arthroplasty:

- The methodology and/or results of systematic item development. The salient steps for development of outcome instruments included: the identification of relevant items from existing studies evaluating the methodological quality of outcome instruments, identification of all patient-, knee-, limb-, defect-, and intervention-specific parameters relevant to patients undergoing articular cartilage surgery of the knee, identification of all measures of outcome assessment, item generation, item reduction, instrument pre-testing, and item weighting.

- The validity of outcome measures, including: content validity, representing the ability of an instrument to measure the traits of interest, including all relevant facets of the investigated condition according to subject matter experts; criterion validity, a comparison of the new instrument to a previously established gold standard; and construct validity, or the validity of inferences that observations or measurement tools actually represent or measure the construct being investigated.

- The reliability (inter-observer and intra-observer) of an outcome instrument, meaning its ability to provide the same measurement of outcome, independent of who administers it, or where it is administered.

- The responsiveness of an outcome instrument, or its ability to accurately detect change when it has occurred.

Exclusion criteria. Articles excluded from review included those in which:

- The text did not report data relevant to patient-reported measures of the outcome of TKA.

- The reported outcome measures were not patient-generated.

- The authors did not report at least one psychometric property (development, validity, reliability, responsiveness) of at least one patient-reported outcome score.

- The study only reported the outcomes of surgery, rather than the intrinsic properties of the outcome measure employed.

- The study reported outcomes for a surgical procedure other than TKA, including unicompartmental or bicompartamental knee arthroplasty, tibial or femoral osteotomy, arthroscopy, cartilage restoration, or total hip arthroplasty. In addition, we excluded articles not written in English.

Statistical analysis. Articles meeting these criteria were assessed and the following data were extracted: study design, subject demographics, surgical technique, and follow-up details. After narrowing down the most frequently used PROMS by setting the frequency arbitrarily at five, the resulting instruments were further analysed by the individual items within each instrument. When applicable, descriptive statistics were calculated for each parameter/variable evaluated. Continuous variable data were reported as mean and standard deviation (SD) (weighted means where applicable). Categorical variable data were reported as frequencies with percentages.

Results

Descriptive summary of the study. Table 1 summarises the key elements from the assessed studies. A total of 38 articles were identified for inclusion, which reported outcomes derived from a total of 85 541 subjects.²²⁻⁵⁹ Of these articles, 28^{23,25-27,30,32-34,36-43,46-51,53-58} reported the results of pre-operative surveys and 30^{22-27,30,32-34,36-43,46-51,53-58} reported post-operative surveys. A total of 47 different PROMS were administered at time periods ranging from three to 118 months post-operatively, at a mean of 26.3 months (SD 30.8). Of the 38 studies, eight developed new scores, 20 evaluated existing scores, and ten represented studies analysing and adapting instruments cross-culturally.

Development, validity, reliability, and responsiveness. With respect to the 'gold standard' criteria of development, validity, reliability, and responsiveness, six of the 38 studies fulfilled all criteria.^{26,28,32,37,50,51}

Systematic development fulfilling the stringent criteria was present in 21 of 38 studies.^{24,26-29,31,32,34,36-38,42,45,46,48-52,54,56} Of the 21 that were properly developed, seven developed new scores:^{24,32,37,42,48,49,52} FJS, PerF/PerP, KKS, SAPS, New KSS, KOOS-PS, HAAS. Of the 17 studies that were not considered properly developed, no explicit reference to the inclusion criteria was present. Among those that did mention the development of the questionnaire, often absent was a reference to item reduction and instrument pre-testing.

The validity of the outcomes instrument was addressed in 33 of the 38 studies, all of which declared the PROMS to be valid. Five studies did not mention or attempt to examine the validity of the outcome measurements.^{22,38,40,41,57} Among the 33 studies addressing validity, the majority focused upon construct, content and

Table I. Summary of key elements of assessed studies

Total number of studies	38
Studies declaring Level I evidence*	1
Studies declaring Level II evidence*	3
Studies declaring Level III evidence*	5
Single-centre studies	31
Multicentre studies	7
North American studies	5
European studies	27
Asian studies	7
Australian/New Zealand studies	1
Studies declaring conflict of interest	3
Studies declaring no conflict of interest	35
Subjects in 38 studies†	7836
Male subjects (n, %)	2899 (37)
Female subjects (n, %)	4937 (63)
Right Knee TKA count	608
Left Knee TKA count	611
Bilateral TKA count	24
Studies conducting pre-op surveys (n, %)	28 (74)
Studies conducting post-op surveys (n, %)	30 (79)
Studies assessing development (n, %)	21 (55)
Studies assessing validity (n, %)	33 (87)
Studies assessing reliability (n, %)	28 (74)
Studies assessing responsiveness (n, %)	12 (32)
OKS studies	14
KSS studies	15
KOOS studies	6
WOMAC studies	12

* Only 9 out of 38 studies self-declared the level of evidence

†One study was excluded in this value because it did not provide gender breakdown though it assessed the predominant number of subjects (n = 77 705) TKA, total knee arthroplasty; pre-op, pre-operative; post-op, post-operative; OKS, Oxford Knee Score; KSS, Knee Society Score; KOOS, Knee injury and Osteoarthritis Outcome Score; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index

convergence validity, specifically, while some studies addressed the subject of validity more generically. For example, Frie et al³¹ used the odds ratio as a metric of validity, while De Groot et al²⁸ broadly stated that the Dutch version of the Knee injury and Osteoarthritis Outcome Score (KOOS) was valid but also recommended that more studies were needed in order to assess its construct validity.

In total, 28 of the 38 studies^{22-24,26,28,29,31-38,40-43,46-51,54-56,58,59} (74%) examined the reliability (inter- and intra-observer) of outcome instruments and concluded that each instrument was reliable based on the measurements undertaken. All of these addressed intra-observer reliability including test-retest reliability and internal consistency using Cronbach's alpha.

Among the reviewed 38 studies, only 12^{26,28,32,33,37,39,44,45,50,51,59} addressed the responsiveness (delta or absolute) of the survey results, with one⁴⁰ focusing on it solely. A total of 26 studies did not address responsiveness. All the PROMS addressed in the 12 studies were declared to be responsive, even though one⁴⁰ called for a larger sample size to evaluate the responsiveness of the KOOS subscales.

In-depth analysis of PROMS. After narrowing it down to the most frequently used PROMS, four knee-specific instruments remained: The Oxford Knee Score^{60,61} (OKS), The New Knee Society Score⁴⁹ (KSS; 2011 version), The KOOS,⁶² and The Western Ontario and McMaster Universities Osteoarthritis Index⁶³ (WOMAC). Six versions of the OKS were analysed in addition to the original English version^{30,33,35,39,40,55} (Swedish,²⁸ Korean,²⁹ Portuguese,³¹ French,^{34,36} Singaporean,^{38,56} and German⁴⁶). Similarly, the KSS was analysed in English,^{22,23,34-37,39-41,45,46,54} Dutch,⁵⁴ and French,⁵⁹ the KOOS was analysed in English,^{26,44,49-51} and Dutch,²⁷ and the WOMAC was only analysed in English.^{23,24,28,32,33,42,45,46,49-51,53} All adaptations were properly developed and tested on international patients (Table II).

With ten and 12 items, respectively, the KSS and OKS PROMS are shorter than the KOOS and WOMAC. While the OKS is entirely patient-derived, the KSS consists of a PROMS component and an informational component for completion by surgeon and patient. Only the PROMS component of the KSS is used to generate a psychometrically-valid knee score. The KOOS and the WOMAC were longer instruments consisting of 42 and 33 items, respectively.

Table II. Summary of key elements of cross-cultural patient-reported outcome measures (PROMS)

Article / cohort	PROM	Translation method	Psychometric properties
Debette et al ⁵⁹	French KSS	Adapted using guidelines	Valid, reliable, responsive
Van der Straeten et al ⁵⁴	New Dutch KSS	Adapted using guidelines	Valid, reliable, developed
de Groot et al ²⁷	Dutch KOOS, SF-36, VAS	Adapted using guidelines	Valid, developed
Dunbar et al ²⁸	Swedish OKS, WOMAC, SF-12, SIP, SF-36	Adapted using guidelines	Valid, reliable, responsive, developed
Eun et al ²⁹	Korean OKS, SF-36, VAS	Adapted using guidelines	Valid, reliable, developed
Gonçalves et al ³¹	Portuguese OKS, SF-36, VAS	Adapted using guidelines	Valid, reliable, developed
Jenny et al ³⁴	French OKS	Translated	Valid, reliable, developed
Jenny et al ³⁵	French OKS, HAAS	Translated	Valid, reliable
Ko et al ³⁸	Singaporean OKS, SF-36	Adapted using guidelines	Valid, reliable, developed
Xie et al ⁵⁶	OKS Singaporean	Adapted using guidelines	Valid, reliable, developed
Naal et al ⁴⁶	OKS German, WOMAC, KSS, SF-12	Adapted using guidelines	Valid, reliable, developed

KOOS, Knee Injury and Osteoarthritis Outcome Score; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index; KSS, Knee Society Score; OKS, Oxford Knee Score; VAS, visual analogue scale; HAAS, High-Activity Arthroplasty Score; SF-, short form; SIP, sickness impact profile

While both reported symptoms, stiffness, pain, and ADLs, the KOOS lengthened the survey to place emphasis on sports/recreation and quality of life. Similar to the OKS, these instruments are entirely focused on the patient.

Of these four most used PROMS, all broadly assessed pain and activity function. The OKS and KSS provide the patient with one item to capture the patient's general level of pain, while the KOOS and WOMAC dedicated an entire section of nine identical items about pain experienced during specific activities. The OKS does solicit further pain information in the context of general activities (i.e. work interference and walking), whereas the KOOS and WOMAC force the patient to recall specific knee manoeuvres (i.e. knee twisting and straightening). In terms of function, the only common function across all four PROMS was the patient's comfort level handling stair ascent and descent. The OKS, KOOS, and WOMAC assessed the knee with respect to kneeling, transportation, capacity for domestic work (household shopping, cooking, etc.), and bathroom activities (ability to wash and dry, getting on and off the toilet, and so on). The OKS and KSS shared only the item the distance the patient can walk. Alternatively, the KSS shared the objective portion of flexion contracture and extension lag, with items within the symptoms sections of both KOOS and WOMAC asking for the patient's ability to bend fully and extend the knee. Among all four PROMS, the KOOS uniquely accounted for ADLs related to sports and recreation. Furthermore, it was the only PROM that asked quality of life items.

It is to be noted that the only alternate form of the four PROMS that was inherently distinct and not simply a cross-cultural translation of the original was the KSS which included both patient-reported, and surgeon-reported, components. There are four separate domains (clinical, functional, satisfaction, and fulfilment of expectation), with the clinical domain being a surgeon-completed section not used as part of the numerical score. Unique to this PROM is its query of both high-demand activities and three priority activities that are self-selected

from a list of popular activities most important to the patient.

Discussion

The value of healthcare interventions has come under recent scrutiny in order to provide both high-quality and cost-effective care. This has resulted in efforts to capture the patient's own evaluation of the outcome of their care using instruments that attempt to measure an amalgamation of attributes of the ideal operation, including relief of pain, restoration of function, and fulfilment of patients' expectations. Hence, the advent of PROMs and their seeming ubiquity in contemporary orthopaedics has led to dozens of different 'outcome instruments' that purport to measure numerous patient-reported and clinician-measured outcomes following TKA.

This systematic review examined 38 studies, in which 47 different PRO instruments were administered to patients following TKA. Of the 38 studies, only six assessed all four of the 'gold standard' criteria, two of which introduced new scores,^{32,37} of which neither were the commonly studied scores (KSS, WOMAC, KOOS, OKS). In all, 33 publications assessed the validity of one or more instruments, 28 the reliability, 12 the responsiveness, and 21 the methodology of initial development. Given the absence of analysis meeting all four 'gold standard' criteria of the most common scores, our review established that a single, properly developed, validated, reliable, and responsive PROM is lacking. In addition, significant inconsistencies were observed between studies in the administration and use of the instruments. Only 28 of the 38 studies (74%) reported pre-operative survey measures, and 30 studies reported post-operative survey measures (79%). The time at which each instrument was administered, post-operatively, varied extensively between instruments and studies ranging from three to 118 months (mean of 26.3 months; SD 30.8), which limits the validity of comparisons between studies, especially those including data collected at less than 12 months post-surgery. Determining the completion rate of many

of the PROMS surveyed and the education level of the respondents was not possible due to inconsistent reporting or inextricability. The significant heterogeneity in administration of the PROMS imparts an additional layer of irregularity in achieving comparable outcome metrics.

Many of the PROMS reviewed grouped items within two domains (pain and function). Some PROMS addressed ADLs (including sports and recreation activities) and quality of life measures. The most complex instrument (the SF-36) addressed eight domains (physical functioning, physical role, bodily pain, general health, vitality, social functioning, emotional role, and mental health). Some PROMS tailored items to address three to five domains of interest. Though all of the PROMS reviewed attempt to measure the success of TKA from the patient's perspective, few specifically address patient satisfaction.^{2,21,27,36}

The proliferation of questionnaires evaluating patient outcomes after TKA, and the variety of items within these instruments, suggests an absence of consensus on the facets of 'outcome' that should be addressed by measurement tools. However, this impression belies the fact that most scores are highly correlated, in keeping with the established methodology for establishing content validity of new measurement instruments. This suggests that the developers of the available instruments may have initially attempted to include different indicators of outcome, only to find that they lacked unidimensionality or were not adequately correlated with legacy measures. This raises some fundamental questions: should additional dimensions of patient outcome be measured to assess the results of TKA? Are patient-reported (i.e. subjective) items sufficient for quantifying patient outcome or should future outcome scores include both subjective and objective measures?

Despite the similarity of many of the items making up the four most popular PROMS, it is clear that three (OKS, WOMAC, KOOS) were designed solely for patients, whereas the KSS has an 'objective' section primarily designed for completion by surgeons which is reported separately from the PROMS instruments that measure function, satisfaction and fulfilment of expectations. There was significant overlap between PROMS in items assessing both pain and function. While the KOOS and WOMAC offered more ways of answering items regarding pain than the OKS and the KSS, the items regarding pain remained similar. In the area of function, the ability to perform stair ascent and descent were of universal interest. ADLs such as the ability to perform domestic chores (household shopping, cooking) and bathroom activities (ability to wash and dry) were examined in three of the four PROMS. The KOOS was unique in asking about high-demand activities and global quality of life, which recognises the patient's desire to perform the high-demand activities that fully restore quality of life through regained confidence in the function of the knee following TKA.


No explicit definition of a successful outcome was discovered among the 38 studies analysing the PRO instruments. Moreover, this systematic review confirms that no single instrument exists to measure the outcome of this procedure. An optimal PROM would encompass all aspects of outcome valued by each patient and perceived as important by the healthcare provider. The instrument would broadly measure a wide range of patient demographics and expectations. However, the ideal instrument should also be specific enough to minimise the impact of comorbidities affecting the general health of the patient in distinction to the function of the knee itself. In deciding which PROM instrument is best suited for measuring TKA functional status, all four instruments were found to be unidimensional in soliciting the same queries (pain, knee function, daily function). Of the four PROMS instruments identified by this systematic review, the KSS 2011 provides the greatest accommodation of variations between the lifestyles of individual patients as it includes both advanced activities and three self-selected activities of greatest importance to the patient.

The ability of any systematic review to answer the important questions sought from the literature is fundamentally limited by the quality of the studies that are available for review. We limited the scope of this review to studies that evaluated the psychometric properties of PROMS used to measure the outcome of TKA. Performance and detection bias was present across patient-specific (completion of questionnaires), clinician-specific (administration, recording, and reporting of patient responses), and instrument-specific (number and variety of items) variables. Among the PROMS evaluated in this review, a primary limitation was the absence of documentation of the patient's primary ADL, the patient's pre-operative condition, and the surgeon's prognostic evaluation of the benefits of TKA relative to those pre-existing conditions. Consequently, the reviewed PROMS did not measure the convergence of the patient's post-TKA expectations with the surgeon's prognosis which is a key factor in obtaining patient consent for a TKA procedure.

Available new information should be developed into a globally shared database and sorted by the key influencing parameters, to eventually yield reliable predictive models. Until such a measure is available, the psychometric validity of the existing PROMS instruments will continue to be debated within the context of their relevance to each patient's individual definition of a successful outcome after TKA. Of the four PROMS instruments identified by this systematic review, the KSS 2011 provides the greatest accommodation of variations between the lifestyles of individual patients as it includes both advanced activities and three self-selected activities of greatest importance to the patient. This approach has the potential of facilitating shared decision making by defining the patient's symptoms, aspirations and expectations. This may form a basis for reconciling the patient's desires with

the clinician's estimation of the likely outcome of treatment option based on the pre-operative evaluation of the patient and prior clinical experience.

Supplementary material

 Two tables explaining the patient-reported outcome measures (PROMs) addressed in this study, as well as the psychometric properties reported in each study, are available alongside the online version of this article at www.bjr.boneandjoint.org.uk

References

- Noble PC, Conditt MA, Cook KF, Mathis KB. The John Insall Award: patient expectations affect satisfaction with total knee arthroplasty. *Clin Orthop Relat Res* 2006;452:35–43.
- Noble PC, Gordon MJ, Weiss JM, et al. Does total knee replacement restore normal knee function? *Clin Orthop Relat Res* 2005;431:157–165.
- Noble PC, Scuderi GR, Brekke AC, et al. Development of a new Knee Society scoring system. *Clin Orthop Relat Res* 2012;470:20–32.
- Alviar MJ, Olver J, Brand C, Hale T, Khan F. Do patient-reported outcome measures used in assessing outcomes in rehabilitation after hip and knee arthroplasty capture issues relevant to patients? Results of a systematic review and ICF linking process. *J Rehabil Med* 2011;43:374–381.
- Collins NJ, Roos EM. Patient-reported outcomes for total hip and knee arthroplasty: commonly used instruments and attributes of a "good" measure. *Clin Geriatr Med* 2012;28:367–394.
- Lingard EA, Sledge CB, Learmonth ID, Kinemax Outcomes Group. Patient expectations regarding total knee arthroplasty: differences among the United States, United Kingdom, and Australia. *J Bone Joint Surg [Am]* 2006;88-A:1201–1207.
- Gandhi R, Davey JR, Mahomed N. Patient expectations predict greater pain relief with joint arthroplasty. *J Arthroplasty* 2009;24:716–721.
- Wylde V, Blom AW. Assessment of outcomes after hip arthroplasty. *Hip Int* 2009;19:1–7.
- Jones CA, Beaupre LA, Johnston DW, Suarez-Almazor ME. Total joint arthroplasties: current concepts of patient outcomes after surgery. *Rheum Dis Clin North Am* 2007;33:71–86.
- Jourdan C, Poiraudou S, Descamps S, et al. Comparison of patient and surgeon expectations of total hip arthroplasty. *PLoS One* 2012;7:30195.
- Bourne RB, Maloney WJ, Wright JG. An AOA critical issue. The outcome of the outcomes movement. *J Bone Joint Surg [Am]* 2004;86-A:633–640.
- Seil R, Pape D. Causes of failure and etiology of painful primary total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc* 2011;19:1418–1432.
- Bourne RB. Measuring tools for functional outcomes in total knee arthroplasty. *Clin Orthop Relat Res* 2008;466:2634–2638.
- Ghomrawi HM, Franco Ferrando N, et al. How Often are Patient and Surgeon Recovery Expectations for Total Joint Arthroplasty Aligned? Results of a Pilot Study. *HSS J* 2011;7:229–234.
- Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KD. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? *Clin Orthop Relat Res* 2010;468:57–63.
- Groeneveld PW, Kwok CK, Mor MK, et al. Racial differences in expectations of joint replacement surgery outcomes. *Arthritis Rheum* 2008;59:730–737.
- Noble PC, Dwyer M, Brekke A. Commonalities, differences, and challenges with patient-derived outcome measurement tools: function/activity scales. *Clin Orthop Relat Res* 2013;471:3457–3465.
- Cordero-Ampuero J, Darder A, Santillana J, Caloto MT, Nocea G. Evaluation of patients' and physicians' expectations and attributes of osteoarthritis treatment using Kanoo methodology. *Qual Life Res* 2012;21:1391–1404.
- Stevens M, Reininga IH, Bulstra SK, Wagenmakers R, van den Akker-Scheek I. Physical activity participation among patients after total hip and knee arthroplasty. *Clin Geriatr Med* 2012;28:509–520.
- Noble PC, Fuller-Lafreniere S, Meftah M, Dwyer MK. Challenges in outcome measurement: discrepancies between patient and provider definitions of success. *Clin Orthop Relat Res* 2013;471:3437–3445.
- Howick J, Chalmers I, Glasziou P, et al. Explanation of the 2011 Oxford Centre for Evidence-Based Medicine (OCEBM) Levels of Evidence (Background Document). Oxford Centre for Evidence-Based Medicine. <http://www.cebm.net/index.aspx?o=5653> (date last accessed 10 February 2015).
- Bach CM, Nogler M, Steingruber IE, et al. Scoring systems in total knee arthroplasty. *Clin Orthop Relat Res* 2002;399:184–196.
- Becker R, Döring C, Denecke A, Brosz M. Expectation, satisfaction and clinical outcome of patients after total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc* 2011;19:1433–1441.
- Behrend H, Giesinger K, Giesinger JM, Kuster MS. The "forgotten joint" as the ultimate goal in joint arthroplasty: validation of a new patient-reported outcome measure. *J Arthroplasty* 2012;27:430–436.
- Busija L, Osborne RH, Nilsdotter A, Buchbinder R, Roos EM. Magnitude and meaningfulness of change in SF-36 scores in four types of orthopedic surgery. *Health Qual Life Outcomes* 2008;6:55.
- Davis AM, Perruccio AV, Canizares M, et al. Comparative, validity and responsiveness of the HOOS-PS and KOOS-PS to the WOMAC physical function subscale in total joint replacement for osteoarthritis. *Osteoarthritis Cartilage* 2009;17:843–847.
- de Groot IB, Favejee MM, Reijman M, Verhaar JAN, Terwee CB. The Dutch version of the knee injury and osteoarthritis outcome score: A validation study. *Health Qual Life Outcomes* 2008;6:16.
- Dunbar MJ, Robertsson O, Ryd L, Lidgren L. Translation and validation of the Oxford-12 item knee score for use in Sweden. *Acta Orthop Scand* 2000;71:268–274.
- Eun IS, Kim OG, Kim CK, Lee HS, Lee JS. Validation of the Korean version of the Oxford Knee Score in patients undergoing total knee arthroplasty. *Clin Orthop Relat Res* 2013;471:600–605.
- Grosse Frie K, van der Meulen J, Black N. Relationship between patients' reports of complications and symptoms, disability and quality of life after surgery. *Br J Surg* 2012;99:1156–1163.
- Gonçalves RS, Tomás AM, Martins DI. Cross-cultural adaptation and validation of the Portuguese version of the Oxford Knee Score (OKS). *Knee* 2012;19:344–347.
- Hossain FS, Patel S, Fernandez MA, Konan S, Haddad FS. A performance based patient outcome score for active patients following total knee arthroplasty. *Osteoarthritis Cartilage* 2013;21:51–59.
- Impellizzeri FM, Mannion AF, Leunig M, Bizzini M, Naal FD. Comparison of the reliability, responsiveness, and construct validity of 4 different questionnaires for evaluating outcomes after total knee arthroplasty. *J Arthroplasty* 2011;26:861–869.
- Jenny JY, Louis P, Diesinger Y. High activity arthroplasty score has a lower ceiling effect than standard scores after knee arthroplasty. *J Arthroplasty* 2014;29:719–721.
- Jenny JY, Diesinger Y. The Oxford Knee Score: compared performance before and after knee replacement. *Orthop Traumatol Surg Res* 2012;98:409–412.
- Jenny JY, Diesinger Y. Validation of a French version of the Oxford knee questionnaire. *Orthop Traumatol Surg Res* 2011;97:267–271.
- 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. *Clin Orthop Relat Res* 2013;471:1539–1547.
- Ko Y, Lo NN, Yeo SJ, et al. Rasch analysis of the Oxford knee score. *Osteoarthritis Cartilage* 2009;17:1163–1169.
- Ko Y, Lo NN, Yeo SJ, et al. Comparison of the responsiveness of the SF-36, the Oxford Knee Score, and the Knee Society Clinical Rating System in patients undergoing total knee replacement. *Qual Life Res* 2013;22:2455–2459.
- Liow RYL, Walker K, Wajid MA, Bedi G, Lennox CME. The reliability of the American Knee Society score. *Acta Orthop Scand* 2000;71:603–608.
- Liow RYL, Walker K, Wajid MA, Bedi G, Lennox CME. Functional rating for knee arthroplasty: comparison of three scoring systems. *Orthopedics* 2003;26:143–149.
- Mahomed N, Gandhi R, Daltroy L, Katz JN. The self-administered patient satisfaction scale for primary hip and knee arthroplasty. *Arthritis* 2011;2011:591253.
- Marx RG, Jones EC, Atwan NC, et al. Measuring improvement following total hip and knee arthroplasty using patient-based measures of outcome. *J Bone Joint Surg [Am]* 2005;87-A:1999–2005.
- Monticone M, Ferrante S, Salvaderi S, Motta L, Cerri C. Responsiveness and minimal important changes for the Knee Injury and Osteoarthritis Outcome Score in subjects undergoing rehabilitation after total knee arthroplasty. *Am J Phys Med Rehabil* 2013;92:864–870.
- Na SE, Ha CW, Lee CH. A new high-flexion knee scoring system to eliminate the ceiling effect. *Clin Orthop Relat Res* 2012;470:584–593.
- Naal FD, Impellizzeri FM, Sieverding M, et al. The 12-item Oxford Knee Score: cross-cultural adaptation into German and assessment of its psychometric properties in patients with osteoarthritis of the knee. *Osteoarthritis Cartilage* 2009;17:49–52.
- Naal FD, Impellizzeri FM, Leunig M. Which is the best activity rating scale for patients undergoing total joint arthroplasty? *Clin Orthop Relat Res* 2009;467:958–965.
- Noble PC, Scuderi GR, Brekke AC, et al. Development of a new Knee Society scoring system. *Clin Orthop Relat Res* 2012;470:20–32.
- Perruccio AV, Stefan Lohmander L, Canizares M, et al. The development of a short measure of physical function for the knee OA KOOS-Physical Function Short-form (KOOS-PS) - an OARSI/OMERACT initiative. *Osteoarthritis Cartilage* 2008;16:542–550.

50. **Roos EM, Toksvig-Larsen S.** Knee injury and Osteoarthritis Outcome Score (KOOS) – validation and comparison to the WOMAC in total knee replacement. *Health Qual Life Outcomes* 2003;1:17.
51. **Ruyssen-Witrand A, Fernandez-Lopez CJ, Gossec L, et al.** Psychometric properties of the OARSI/OMERACT osteoarthritis pain and functional impairment scales: ICOAP, KOOS-PS and HOOS-PS. *Clin Exp Rheumatol* 2011;29:231–237.
52. **Talbot S, Hooper G, Stokes A, Zordan R.** Use of a new high-activity arthroplasty score to assess function of young patients with total hip or knee arthroplasty. *J Arthroplasty* 2010;25:268–273.
53. **Toyé F, Barlow J, Wright C, Lamb SE.** A validation study of the New Zealand score for hip and knee surgery. *Clin Orthop Relat Res* 2007;464:190–195.
54. **Van Der Straeten C, Witvrouw E, Willems T, Bellemans J, Victor J.** Translation and validation of the Dutch New Knee Society scoring system. *Clin Orthop Relat Res* 2013;471:3565–3571.
55. **Whitehouse SL, Blom AW, Taylor AH, Pattison GTR, Bannister GC.** The Oxford Knee Score; problems and pitfalls. *Knee* 2005;12:287–291.
56. **Xie F, Li SC, Lo NN, et al.** Cross-cultural adaptation and validation of Singapore English and Chinese Versions of the Oxford Knee Score (OKS) in knee osteoarthritis patients undergoing total knee replacement. *Osteoarthritis Cartilage* 2007;15:1019–1024.
57. **Zhang XH, Li SC, Xie F, et al.** An exploratory study of response shift in health-related quality of life and utility assessment among patients with osteoarthritis undergoing total knee replacement surgery in a tertiary hospital in Singapore. *Value Health* 2012;15 (suppl):S72–S78.
58. **Dinjens RN, Senden R, Heyligers IC, Grimm B.** Clinimetric quality of the new 2011 Knee Society score: high validity, low completion rate. *Knee* 2014;21:647–654.
59. **Debette C, Parratte S, Maucort-Boulch D, et al.** French adaptation of the new Knee Society Scoring System for total knee arthroplasty. *Orthop Traumatol Surg Res* 2014;100:531–534.
60. **Dawson J, Fitzpatrick R, Murray D, Carr A.** Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg [Br]* 1998;80-B:63–69.
61. **Murray DW, Fitzpatrick R, Rogers K, et al.** The use of the Oxford hip and knee scores. *J Bone Joint Surg [Br]* 2007;89-B:1010–1014.
62. **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.
63. **Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW.** Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;15:1833–1840.

Funding statement:

- None declared

Author contributions:

- P. N. Ramkumar: Performed the search, analysis and writing of the manuscript
- J. D. Harris: Performed the analysis and editing of the manuscript
- P. C. Noble: Performed the analysis, writing and editing of the manuscript

ICMJE Conflict of Interest:

- None declared

©2015 Noble et al. This is an open-access article distributed under the terms of the Creative Commons Attributions licence, which permits unrestricted use, distribution, and reproduction in any medium, but not for commercial gain, provided the original author and source are credited.