Clinical Orthopaedics
and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

CrossMark

CLINICAL RESEARCH

Validation of the KOOS, JR: A Short-form Knee Arthroplasty Outcomes Survey

Stephen Lyman PhD, Yuo-Yu Lee MS, Patricia D. Franklin MD, MBA, MPH,
Wenjun Li PhD, Michael B. Cross MD, Douglas E. Padgett MD

Abstract

Background Medicare is rapidly moving toward using patient-reported outcome measures (PROMs) for outcomes assessment and justification of orthopaedic and other procedures. Numerous measures have been developed to study knee osteoarthritis (OA); however, many of these surveys are long, disruptive to clinic flow, and result in incomplete data capture and/or low followup rates. The Knee injury and Osteoarthritis Outcome (KOOS) physical function short-form (KOOS-PS), while shorter, ignores pain, which is a primary concern of patients with advanced knee OA.

Questions/purposes Our objective was to derive and validate a short-form survey focused on the patient with end-stage knee OA undergoing TKA.

Methods Using our hospital's knee replacement registry, we retrospectively identified 2291 patients with knee OA who underwent primary unilateral TKA and had completed preoperative and 2-year postoperative PROMs. We assessed 30 items from the 42-item KOOS that were quantitatively most difficult for patients to perform before TKA and qualitatively most relevant to patients with end-stage knee OA. Rasch analysis identified the KOOS, JR, a seven-item instrument, representing a single dimension, which we define as "knee health" because it reflects aspects of pain, symptom severity, and activities of daily living (ADL) including movements or activities that are directly relevant and difficult for patients with advanced knee OA. We assessed the internal consistency, external

S. Lyman (✉), Y.-Y. Lee
Healthcare Research Institute, Hospital for Special Surgery, 530 East 70th Street, New York, NY 10021, USA
e-mail: lymans@hss.edu

P. D. Franklin, W. Li
Department of Orthopedics and Physical Rehabilitation, University of Massachusetts Medical School, Worcester, MA, USA

M. B. Cross, D. E. Padgett
Adult Reconstruction and Joint Replacement Service, Hospital for Special Surgery, New York, NY, USA

Springer

validity (versus KOOS and WOMAC domains), responsiveness, and floor and ceiling effects of the KOOS, JR. External validation was performed using calculated KOOS, JR scores in collaboration with a nationally representative joint replacement registry, the Function and Outcomes Research for Comparative Effectiveness in Total Joint Replacement (FORCE-TJR).

Results Internal consistency for the KOOS, JR was high (Person Separation Index, 0.84; and 0.85 [FORCE]), external validity against other validated knee surveys was excellent (Spearman correlation coefficient, ρ 0.54–0.91), particularly for the KOOS pain (ρ 0.89 [95% CI, 0.88–0.91] Hospital for Special Surgery [HSS]; and 0.91 [95% CI, 0.90–0.93] [FORCE]) and KOOS ADL (ρ 0.87 [95% CI, 0.85–0.88] [HSS]; and 0.84 [95% CI, 0.81–0.87] [FORCE]). The KOOS, JR responsiveness (standardized response means, 1.79 [95% CI, 1.70–1.88] [HSS]; and 1.70 [95% CI, 1.54–1.86] [FORCE]) was high and floor 0.4–1.2%) and ceiling (18.8–21.8%) effects were favorable.

Conclusions The new short knee PROM, the KOOS, JR, provides a single score representing "knee health" as it combines pain, symptoms, and functional limitations in a single score. This short-form PROM is patient-relevant and efficient.

Level of Evidence   Level III, diagnostic study.

Introduction

The TKA is the leading cost driver for Medicare, and with the number of TKAs expected to skyrocket between now and 2030 by more than 400% to 3.48 million annually, efficient assessment of TKA outcomes will be increasingly important for patients, clinicians, and payers [11]. The Centers for Medicare & Medicaid Services (CMS) recently released their proposed TKA outcomes assessment rules to meet their pay-for-performance measures [5]. Various validated patient-reported outcome measures (PROMs) for the knee, including the Knee Injury Osteoarthritis Outcomes Survey (KOOS), Oxford Knee Score, and WOMAC, have been used to assess the outcomes of TKA and other treatments for end-stage knee osteoarthritis (OA) [3, 4, 6, 10, 16, 17, 19], but only the KOOS was adopted by the CMS.

Unfortunately, most PROMs are lengthy (KOOS, WOMAC), making widespread administration incompatible with efficient clinic flow and presenting challenges for followup with nonresponse and missing responses at issue. Further, some PROMs are proprietary (WOMAC, Oxford), making them incompatible with the CMS PROM requirements. A shorter, yet well-validated, survey would create efficiencies and potentially improve TKA outcomes

assessments for various purposes including pay-for-performance assessment. The KOOS-PS, which is a short physical function survey derived from the full KOOS, was developed in a more-athletic population and explicitly excluded pain items, which are directly relevant to patients with OA [17]. We therefore used the full KOOS to develop the KOOS Joint Replacement (KOOS, JR), a short-form survey retaining the most relevant questions for patients with end-stage knee OA who are undergoing TKA; the tool we developed seeks to represent a single domain (unidimensional construct) of "knee health", combining pain, symptoms, and functional ability into a single score that is more efficient to administer while still providing an accurate assessment of the patient's knee.

We sought to validate that short-form survey in terms of internal consistency, external validity (versus KOOS and WOMAC domains), responsiveness, and floor and ceiling effects.

Patients and Methods

In designing this study we considered the outcomes measure criteria recommend by Fitzpatrick et al. [8] of appropriateness, reliability, validity, responsiveness, precision, interpretability, acceptability, and feasibility. Because we endeavored to derive a short-form PROM rather than develop a new instrument, we relied on the framework proposed by Rothman et al. [20] in the 2009 International Society for Pharmacoeconomics and Outcomes Research (ISPOR) report on use and modification of existing PROMs.

Subjects

Through our institutional review board-approved Total Knee Replacement Registry, we retrospectively reviewed all patients who underwent primary unilateral TKA for OA at the Hospital for Special Surgery (HSS) between May 1, 2007, and February 28, 2012. Our institutional registry prospectively collected patient demographics and PROMs for TKA, including the KOOS and WOMAC.

Approximately 84% of patients undergoing unilateral primary TKA for OA consented to registry participation, at which time they were administered a preoperative KOOS survey. Approximately 81% of these patients returned a baseline survey. Of these, approximately 79% also returned a 2-year survey. Patients who returned a baseline and 2-year followup survey were eligible for inclusion in the KOOS, JR item eligibility assessment (n = 3772). However, as required by the Rasch model, only patients who completed all items eligible for the formal item reduction

process at baseline and 2 years were included in the item reduction and validation phases (n = 2291; mean age, 67 years, 57% female). The large decrease in eligible patients was the result of several frequently skipped items being deemed eligible for inclusion. For example, "difficulty squatting" was eligible, but 19% of patients skipped the item, which eliminated all patients who did not complete the squatting item. We detected differences between responders and nonresponders in the HSS cohort (data not shown), however, these differences were generally small and unlikely to compromise the validity of this study's conclusions.

Using a computerized randomization algorithm, this cohort was randomly divided into learning (n = 1146) and validation (n = 1145) cohorts for the purpose of building (learning) and validating (validation) the new PROM. External validation was performed using the Function and Outcomes Research for Comparative Effectiveness in Total Joint Replacement (FORCE-TJR) registry (n = 1114), a nationally representative sample of patients who underwent TKAs from across the United States [2, 9].

Item Eligibility Assessment

The KOOS consists of 42 self-administered items (Table 1) in five domains: pain (10 items); symptoms (seven items); limitations in activities of daily living (ADL) (17 items); sports and recreation (four items); and knee-related quality of life (QOL) (four items). We omitted the items from the QOL domain because unlike other KOOS items, they do not address specific knee movements or activities. A preliminary feasibility analysis for this study using Rasch analysis also resulted in the exclusion of all QOL questions, suggesting this was an appropriate decision.

Before initiating the validation, we queried 30 patients, before undergoing TKA, from four surgeon's offices to rate each item in the KOOS on a scale from 1 to 3 with respect to their importance (1 = unimportant, 2 = somewhat important, 3 = very important). These patients were not statistically different from the full cohort used for item reduction and validation (Table 2). Thirty patients were chosen based on previous studies of validity pretesting [1, 20]. Based on these ratings, we calculated mean relevance scores for each item. Thresholds for inclusion were the same as those previously used in a validation of the Foot and Ankle Outcome Score [18]. Specifically, we considered items to be eligible for inclusion in the KOOS, JR if they had a mean relevance score of at least 2.0 or greater with at least ⅔ (67%) of patients rating the item as at least "somewhat relevant." After selecting the relevant questions for validation, we calculated mean scores at baseline for all items. Items with a mean score less than 1.8 were

considered items without sufficient difficulty and were excluded from KOOS, JR eligibility. A score of 2.0 is considered "moderate" difficulty. This lower threshold allows for some margin of error.

After removing items without sufficient difficulty, we further removed redundant items querying the same activity in different domains of the KOOS. These parameters, including twisting/pivoting on your injured knee, ascending or descending stairs, walking on a flat or uneven surface, and standing upright, appear in the domains for pain and ADL or sports and recreation. For example, patients were asked about "pain when twisting/pivoting" and "difficulty when twisting/pivoting". Rather than include both items in the Rasch analysis, we assessed only the item considered more relevant and difficult. Based on the relevance and difficulty calculations described, we found that the items querying pain regarding these activities were more relevant and difficult, therefore we excluded the redundant ADL items. In the case of the sports and recreation item "difficulty when twisting/pivoting", the sports and recreation domain item was slightly more relevant and difficult, but was skipped more than twice as often (9% v. 19%) and was less responsive than the analogous pain item, and was therefore excluded from further assessment. We also removed items with greater than 20% of values missing: running (28% missing) and jumping (30% missing). We believe that the percent missing is high because patients with severe knee OA often skip items that are irrelevant to their daily lives. This left 19 items from the original KOOS for formal item reduction assessment.

Statistical Analysis

The statistical approaches used in this study were described in detail in a related work regarding development and validation of the HOOS, JR [14].

Item Reduction Process

Item reduction was done using a Rasch analysis using a partial credit model [15]. The most basic form of the Rasch model is based on a binary response scale. The partial credit model is an extension to the basic Rasch model and is devised for responses in which one has two or more ordered categories. It permits each item to have its own unique number of categories and modeled distance between adjacent categories. To refine the most likely candidate items for exclusion, bootstrapping of 500 samples of 1800 patients was performed with replacement so patients could be selected to each sample more than once. Bootstrapping is a resampling technique to estimate the

Table 1. Results from relevance survey of KOOS items

KOOS items and subdomains	Importance (n = 30) Mean ± SD	Percent missing from baseline (n = 2291)	Mean difficulty (baseline) Mean ± SD	Standardized response means (2 years)	500 bootstrapping results Percent retained
Symptoms					
S1. Do you have swelling in your knee?	2.66 ± 0.65	4.0	2.28 ± 1.30	0.951	0.0
S2. Do you feel grinding, hear clicking, or any other type of noise when your knee moves?	2.31 ± 0.90	3.0	2.37 ± 1.23	0.769	0.0
S3. Does your knee catch or hang up when moving?	2.35 ± 0.88	5.0	1.50 ± 1.12	0.875	N/A
S4. Can you straighten your knee fully?	2.50 ± 0.76	4.0	1.54 ± 1.48	0.847	N/A
S5. Can you bend your knee fully?	2.53 ± 0.72	3.0	1.93 ± 1.51	0.762	0.0
S6. How severe is your knee stiffness after first awakening in the morning?	2.56 ± 0.67	1.0	2.12 ± 0.96	1.091	97.0*
S7. How severe is your knee stiffness after sitting, lying, or resting later in the day?	2.69 ± 0.59	1.0	2.17 ± 0.90	1.122	99.0*
Pain					
P1. How often do you experience knee pain?	2.78 ± 0.42	6.0	3.17 ± 0.68	1.750	72.0*
P2. Twisting/pivoting on your knee	2.69 ± 0.59	9.0	2.47 ± 0.98	1.602	79.0*
P3. Straightening knee fully	2.58 ± 0.72	6.0	1.83 ± 1.08	1.294	58.0*
P4. Bending knee fully	2.60 ± 0.62	6.0	2.13 ± 1.10	1.175	3.0
P5. Walking on a flat surface	2.50 ± 0.78	2.0	1.85 ± 0.87	1.475	100.0*
P6. Going up or down stairs	2.90 ± 0.31	3.0	2.56 ± 0.87	1.455	94.0*
P7. At night while in bed	2.47 ± 0.73	3.0	1.51 ± 1.01	0.986	N/A
P8. Sitting or lying	2.37 ± 0.76	3.0	1.30 ± 0.93	0.922	N/A
P9. Standing upright	2.43 ± 0.73	3.0	1.85 ± 0.93	1.370	97.0*
Activities of daily living					
A1. Descending stairs	2.94 ± 0.25	6.0	2.41 ± 0.91	1.344	N/A
A2. Ascending stairs	2.84 ± 0.37	3.0	2.31 ± 0.94	1.345	N/A
A3. Rising from sitting	2.71 ± 0.53	3.0	2.14 ± 0.93	1.255	96.0*
A4. Standing	2.45 ± 0.68	3.0	1.82 ± 0.93	1.357	N/A
A5. Bending to floor/pick up an object	2.39 ± 0.67	2.0	1.90 ± 1.03	1.011	74.0*
A6. Walking on a flat surface	2.52 ± 0.72	2.0	1.75 ± 0.85	1.437	N/A
A7. Getting in/out of car	2.52 ± 0.68	2.0	2.04 ± 0.89	1.225	62.0*
A8. Going shopping	2.19 ± 0.79	4.0	2.04 ± 0.93	1.474	95.0*
A9. Putting on socks/stockings	2.26 ± 0.73	2.0	1.60 ± 1.00	0.923	N/A
A10. Rising from bed	2.55 ± 0.68	2.0	1.74 ± 0.96	1.160	N/A
A11. Taking off socks/stockings	2.30 ± 0.72	3.0	1.54 ± 0.98	0.945	N/A
A12. Lying in bed (turning over, maintaining knee position)	2.53 ± 0.73	2.0	1.62 ± 1.00	1.081	N/A
A13. Getting in/out of bath	2.53 ± 0.68	9.0	1.64 ± 1.04	0.929	N/A
A14. Sitting	2.23 ± 0.77	2.0	1.22 ± 0.94	0.909	N/A

Table 1. continued

KOOS items and subdomains	Importance (n = 30) Mean ± SD	Percent missing from baseline (n = 2291)	Mean difficulty (baseline) Mean ± SD	Standardized response means (2 years)	500 bootstrapping results Percent retained
A15. Getting on/off toilet	2.50 ± 0.73	2.0	1.64 ± 0.97	1.075	N/A
A16. Heavy domestic duties (moving heavy boxes, scrubbing floors, etc)	2.30 ± 0.75	14.0	2.49 ± 1.06	1.163	22.0
A17. Light domestic duties (cooking, dusting, etc)	2.27 ± 0.74	5.0	1.57 ± 0.90	1.189	N/A
Sports and recreation					
SP1. Squatting	2.67 ± 0.61	19.0	2.96 ± 1.07	1.024	0.0
SP2. Running	2.47 ± 0.78	28.0	3.25 ± 1.08	1.001	N/A
SP3. Jumping	2.30 ± 0.84	30.0	3.27 ± 1.09	0.996	N/A
SP4. Twisting/pivoting on your injured knee	2.73 ± 0.52	19.0	2.96 ± 1.04	1.297	N/A
SP5. Kneeling	2.57 ± 0.57	17.0	2.93 ± 1.09	0.788	0.0
Quality of life					
Q1. How often are you aware of your knee problem?	2.90 ± 0.31	5.0	3.54 ± 0.54	1.344	N/A
Q2. Have you modified your lifestyle to avoid activities potentially damaging activities to your knee?	2.90 ± 0.40	4.0	2.90 ± 0.97	1.098	N/A
Q3. How much are you troubled with lack of confidence in your knee?	2.70 ± 0.65	4.0	2.78 ± 0.98	1.493	N/A
Q4. In general, how much difficulty do you have with your knee?	2.87 ± 0.43	4.0	2.80 ± 0.77	1.731	N/A

* More than 50%; KOOS = Knee Disability Osteoarthritis Outcomes Survey; N/A = not applicable for items that were not formally assessed using the bootstrapped Rasch analysis.

Table 2. Demographics of relevancy cohort and full cohort

Demographic	Relevancy cohort (n = 30)*	Full cohort (n = 2291)¶	p value
Age (years)	62 ± 14	67 ± 10	0.09
BMI (kg/m²)	29 ± 5	30 ± 6	0.32
Sex	57% female	57% female	0.95

* For eligibility assessment we interviewed 30 consecutive patients scheduled for primary TKA, from four surgeons; ¶patients who returned baseline and 2-year followup surveys, and completed all items eligible at baseline for the formal item reduction assessment (19 items from the original KOOS) were included in the item reduction and validation.

accuracy of our approximation of all patients using only available patients. The item retention process was automated for the 500 iterative samples. The final Rasch analysis process was performed using the items retained in more than 50% of the 500 models. The entire cohort was split into learning and validation cohorts. The 12 items retained after this iterative process were run through a final Rasch analysis with sequential removal to determine which would remain in the KOOS, JR. Eight items were retained after this process.

In an effort to include only universal activities or movements, we assessed whether the KOOS, JR performed equally well with and without the item "going shopping" because this may represent a culturally dependent activity that is perceived differently between men and women. Because the performance of the KOOS, JR was not affected by removal, this item was dropped, leaving seven items in the final KOOS, JR.

Scoring

The KOOS, JR scoring was on a 100-point scale with 0 representing complete knee disability and 100 representing perfect knee health, just as with the original KOOS. Scoring for the final survey was generated from the final items based on the Rasch-based person score using the validation cohort. A crosswalk table converting raw sum score to the interval level measure scaled from 0 to 100 was provided to facilitate the use and scoring of the KOOS, JR (Appendices 1 and 2. Supplemental material is available with the online version of CORR®.). The KOOS, JR scores were derived from the responses to full KOOS surveys from both registries.

Validation Process

A formal validation process was performed for the KOOS, JR using the validation cohort of a 50% sample of the total

HSS cohort and the external FORCE-TJR registry. The internal consistency of the KOOS, JR instrument was evaluated by a Person Separation Index (PSI), which is similar to Cronbach's alpha. A high PSI, which indicates a strong ability to differentiate between patients with differing ability, is evidence of high internal consistency. A value greater than 0.7 was considered acceptable [7]. Items that are included in Rasch analysis are required to be independent of each other (ie, no appreciable correlation between the items included in the survey). Local independence of the items was confirmed using residual item correlations. Items with residual correlations greater than 0.3 are considered to be locally dependent [23].

To verify the remaining selected items conform to a one-dimensional construct, we used a principal component analysis on the standardized residual. In a successful Rasch analysis, there is only one dimension, called the Rasch dimension, captured by the Rasch model and there should be no presence of subdimensions in the principal component analysis. An eigenvalue of the first factor (the Rasch dimension) greater than three or an eigenvalue of each item greater than 1.4 suggests that additional subdimensions are likely to be present [13, 21]. No items were identified as unacceptably correlated during this process.

To measure responsiveness to treatment (TKA), we used standardized response means (SRM) [12] and compared the KOOS, JR with other validated PROMs (KOOS domains, WOMAC domains) in the validation cohort. Responsiveness was considered high if the SRM was greater than 0.8 [22]. We also calculated floor and ceiling effects for the KOOS, JR and made comparisons with these other validated PROMs. Finally, we assessed external construct validity by comparing the Spearman's correlations between the KOOS, JR and other validated PROMs. We also used a scatterplot overlaying a contour plot based on bivariate kernel density estimation between KOOS, JR and other KOOS domains to visually assess the external correlations. A bandwidth multiplier of one was used for each kernel density estimate. Areas of high density correspond to areas where there are many overlapping points.

Factor analyses were performed using SAS® 9.3 (SAS Institute Inc, Cary, NC, USA) and Rasch analysis using the eRm R Package (The R Foundation, Vienna, Austria).

Results

The HSS cohort included 2291 patients with knee OA, from 48 surgeons, who underwent primary unilateral TKA at HSS between May 2007 and February 2012. Fifty-seven percent of the patients were female, and they had a mean age 67 ± 10 years and mean BMI of 30 ± 6 kg/m². The learning and validation cohorts had similar age, sex, and

BMI distributions. The FORCE-TJR registry cohort consisted of 2668 patients with knee OA, from 128 surgeons across 38 practices from 22 US states, undergoing primary unilateral TKA between June 2011 and January 2013. These patients had a mean age of 67 ± 9 years, a mean BMI of 32 ± 6 kg/m^2, and 63% were female.

Item Reduction

Item reduction yielded a seven-item PROM (KOOS, JR), which retained items from the symptoms, pain, and ADL domains. As noted, this resulted in a one-dimensional survey consisting of seven items that were well fit (Fig. 1).

Validation

The seven-item KOOS, JR survey had high internal consistency (PSI, 0.84 [HSS]; and 0.85 [FORCE]), and all items were confirmed to exist on a single dimension based on principal component analysis on the standardized residuals.

Responsiveness of the KOOS, JR was excellent (SRM, 1.70 [95% CI, 1.54–1.86] [FORCE]; and SRM, 1.79 [95% CI, 1.7–1.88] [HSS]) exceeding the theoretical 0.8 SRM threshold (Fig. 2). The KOOS, JR had the highest SRM of all the surveys considered, except KOOS-pain (SRM, 1.82 [95% CI, 1.63–2.01] [FORCE]; and SRM, 1.97 [95% CI, 1.86–2.08] [HSS]), and was similar to the KOOS-QOL (SRM, 1.68 [95% CI, 1.53–1.83] [FORCE]; and SRM, 1.70 [95% CI, 1.58–1.82] [HSS]).

Construct validity of the KOOS, JR compared with KOOS and WOMAC domains was excellent, as a Spearman's correlation coefficient of 0.8 or greater is considered very high external validity [24]. Correlations with the pain and ADL domains of the KOOS were very high (KOOS-Pain 0.89, [95% CI, 0.88–0.91] [HSS]; 0.91, [95% CI, 0.90–0.93] [FORCE]), (KOOS-ADL 0.87, [95% CI, 0.85–0.88] [HSS]; 0.84, [95% CI, 0.81–0.87] [FORCE]), and those with the pain, stiffness, and function domains of the WOMAC also were high to very high (WOMAC-pain 0.80, [95% CI, 0.77–0.82] [HSS]; 0.82 [95% CI, 0.79–0.86] [FORCE]), (WOMAC-stiffness 0.72, [95% CI, 0.69–0.75] [HSS]; −0.76, [95% CI, 0.72–0.80] [FORCE]), (WOMAC-function 0.87, [95% CI, 0.85-0.88] [HSS]; −0.84, [95% CI, 0.81–0.87] [FORCE]). Correlation for KOOS-symptoms, KOOS-sports and recreation, and KOOS-QOL were moderate (KOOS-symptoms 0.59, [95% CI, 0.55–0.64] [HSS]; 0.69, [95% CI, 0.64–0.74] [FORCE]) (KOOS-sports 0.57, [95% CI, 0.53–0.62] [HSS]; 0.54, [95% CI, 0.47–0.61] [FORCE]); (KOOS-QOL 0.59, [95% CI, 0.54–0.63] [HSS]; 0.58, [95% CI, 0.52–0.64] [FORCE]) (Fig. 3). Floor and ceiling effects also were favorable compared with other instruments (Fig. 4). A kernel density plot was used to visualize the correlation between KOOS, JR and KOOS pain (Fig. 5) and ADL domains (Fig. 6). The KOOS, JR and KOOS pain domains shared a positive correlation at baseline and the change between baseline and 2-year followup. The similar result was found between the KOOS, JR and the KOOS ADL domains at baseline and the change between baseline and 2-year followup.
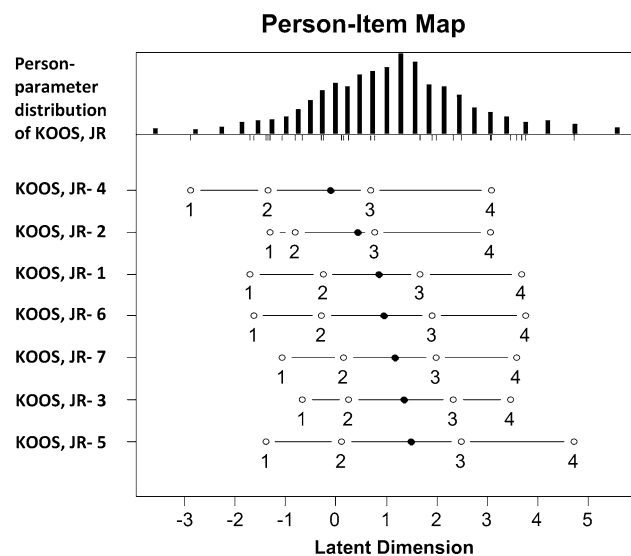

Fig. 1 The person-ability and item difficulty are shown. The horizontal line represents the measure of the variable in linear log units. The top bar graph locates each patient's ability, with ability increasing from right to left. The bottom graph locates each item's relative difficulty for this validation sample, with difficulty increasing from right to left. The numbers represent the thresholds between response categories. For data to adhere to the Rasch model, threshold points are correctly ordered, indicating patients have no difficulty consistently discriminating between response categories. KOOS, JR-1 (Symptom) How severe is your knee joint stiffness after first wakening in the morning?; KOOS, JR- 2 (Pain) Twisting/pivoting on your knee; KOOS, JR- 3 (Pain) Straightening knee fully; KOOS, JR- 4 (Pain) Going up or down stairs; KOOS, JR- 5 (Pain) Standing upright; KOOS, JR- 6 (ADL) Rising from sitting; KOOS, JR- 7 (ADL) Bending to floor/pick up an object.

Discussion

With CMS and other payers rapidly moving toward tying pay-to-performance outcomes, efficient, accurate, nonproprietary PROMs will become increasingly important in assuring compliance with coming regulations. We endeavored to develop a short-form version of the KOOS that was directly relevant to patients undergoing TKA. The KOOS-PS, while short, does not address knee pain, which is a primary concern to patients with knee OA. Therefore,
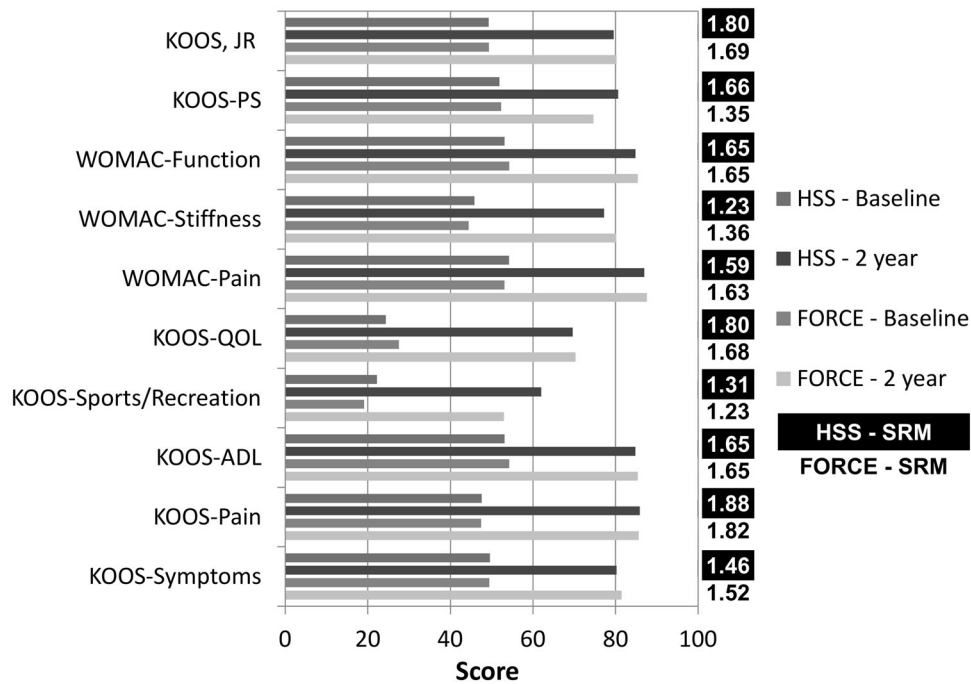
Fig. 2 The standardized response means of knee replacement outcomes measures at preoperative baseline and 2 years after surgery are shown. KOOS-PS = KOOS physical function short-form; QOL = quality of life; ADL = activities of daily living; HSS = Hospital for Special Surgery; FORCE = Function and Outcomes Research for Comparative Effectiveness; SRM = standardized response mean.
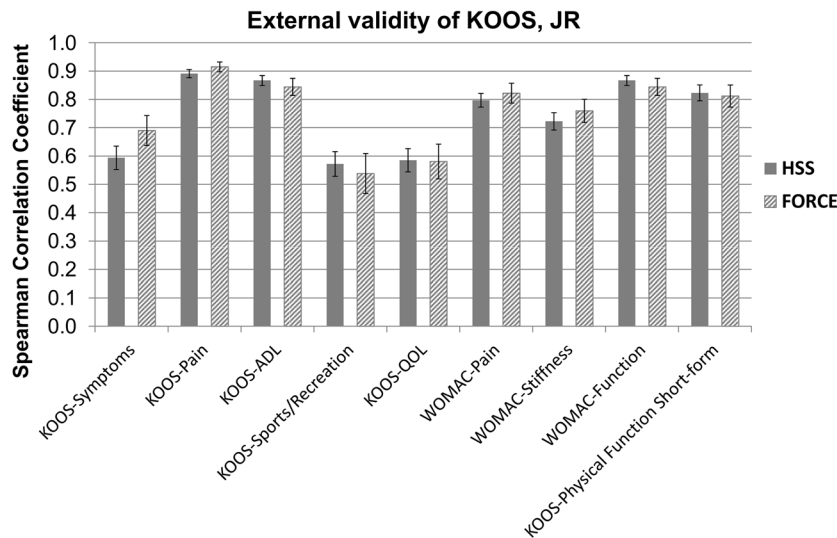

Fig. 3 A comparison of external validity of the KOOS, JR against nine other patient-reported outcome measures using the Spearman correlation coefficient is shown. HSS = Hospital for Special Surgery; FORCE = Function and Outcomes Research for Comparative Effectiveness; ADL = activities of daily living; QOL = quality of life; KOOS-PS = KOOS physical function short-form.

we adapted and validated the KOOS, JR, a novel, seven-question short-form alternative to the longer KOOS and WOMAC surveys for assessing patient-reported outcomes after TKA.

This study has numerous limitations. Our KOOR, JR derivation was performed at one tertiary care musculoskeletal specialty hospital in a dense urban area. The patient population is diverse in socioeconomic status and
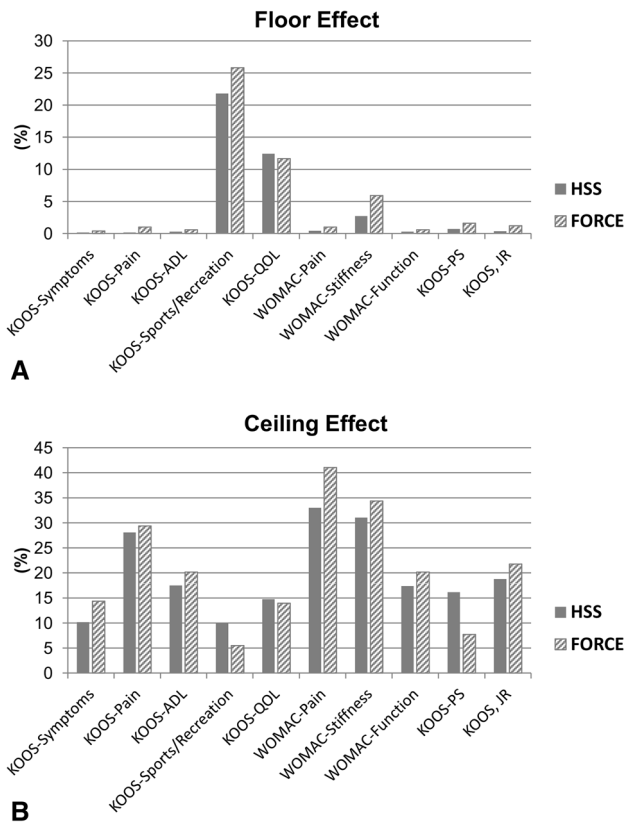
Floor Effect

A

Ceiling Effect

B

Fig. 4A–B (**A**) Floor and (**B**) ceiling effects for 10 patient-reported outcome measures are shown. HSS = Hospital for Special Surgery; FORCE = Function and Outcomes Research for Comparative Effectiveness; ADL = activities of daily living; QOL = quality of life; KOOS-PS = KOOS physical function short-form.

residential environment (including patients drawn from urban, suburban, and rural settings), but unlike the US population at large, most are urban dwellers. However, external validation with the Agency for Healthcare Research and Quality-funded FORCE-TJR registry confirmed the validity of the KOOS, JR for the US population. The KOOS, JR generally performed better in the FORCE-TJR registry population than in the HSS validation cohort population. Similarly, although we detected differences between responders and nonresponders in the HSS cohort (data not shown), in general these differences were small and unlikely to compromise the validity of the study's conclusions, especially since the instrument performed even better in the FORCE-TJR cohort, which was very similar to the nonresponders at HSS.

Another limitation pertained to the retrospective study design. Our study was done as a pragmatic validation process, using existing full KOOS surveys to complete a new short-form survey rather than validating the survey in a new cohort of patients. Patient responses to the full KOOS surveys were used to derive a new short-form survey. No patients were asked to complete the full and the
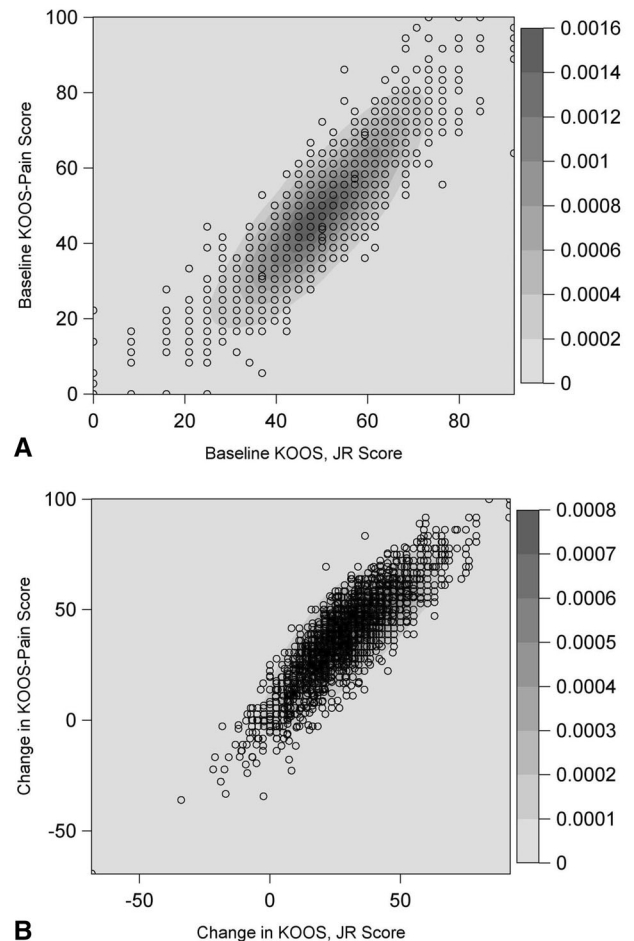

A

B

Fig. 5A–B The contour maps show the KOOS-pain domain versus the (**A**) KOOS, JR at baseline and (**B**) the change in score from baseline to 2 years after THA. A scatterplot overlays a contour plot based on bivariate kernel density estimation. A bandwidth multiplier of one was used for each kernel density estimate. Areas of high density correspond to areas where there are many overlapping points. The scatterplot shows the positive correlation between the KOOS, JR (x-axis) and the KOOS pain domain (y-axis) at baseline and the change between baseline and 2-year followup.

short-form surveys for a direct comparison. Given that item order may influence responses, it is possible that responses to the questions retained for the KOOS, JR would have been answered differently if encountered on their own rather than in the full series of 42 KOOS items. However, this pragmatic process does offer the advantage that, because the validated KOOS, JR was derived from the full KOOS, it can be calculated for all previously administered KOOS surveys for direct comparison.

Finally, the KOOS, JR is limited to relatively low-demand activities (going up and down stairs is the most rigorous activity) (Fig. 1), which may be relevant for older, less-active adults undergoing TKA, but may be inadequate to assess the postoperative experience of younger, more-active patients undergoing TKA. The ceiling effect of
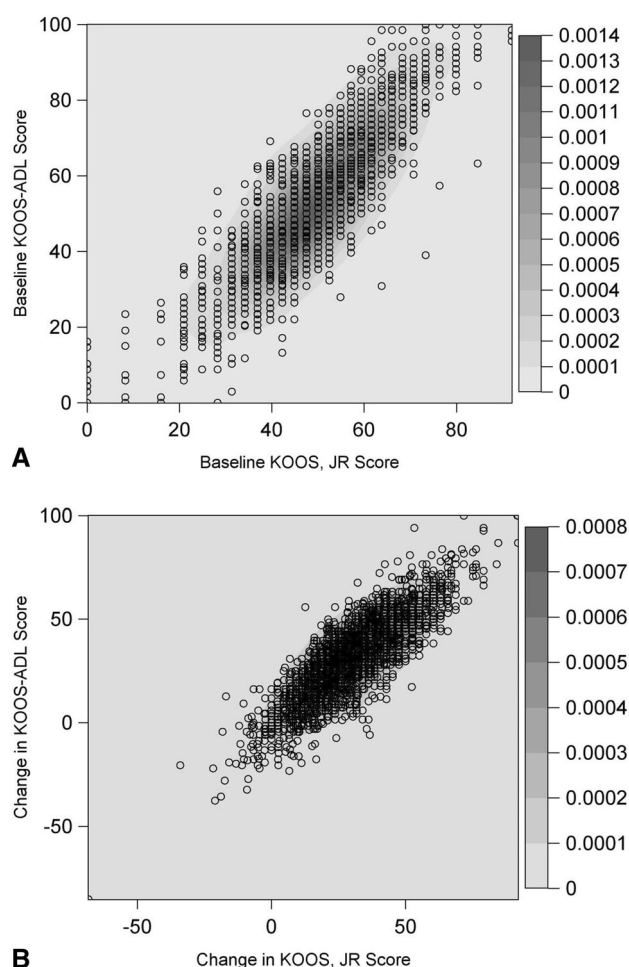
Fig. 6A–B The contour maps show the KOOS-ADL domain versus the (**A**) KOOS, JR at baseline and (**B**) the change in score from baseline to and 2 years after THA. In the figure, a scatterplot overlays a contour plot based on bivariate kernel density estimation. A bandwidth multiplier of one was used for each kernel density estimate. Areas of high density correspond to areas where there are many overlapping points. The scatterplot shows the positive correlation between the KOOS, JR (x-axis) and the KOOS ADL domain (y-axis) at baseline and the change between baseline and 2-year followup. ADL = activities of daily living.

approximately 20% likely reflects this. For these patients it may be advisable to also capture either the KOOS-QOL domain (four items, ceiling below 15%) or KOOS sports and recreation domain (four items, ceiling below 10%).

Validity

Although validity of the KOOS was tested when the survey originally was developed [19], we nevertheless examined specific relevance to patients undergoing TKA and found that all were considered relevant. Our principal component analysis confirmed that the KOOS, JR represents a unidimensional construct; which we defined as "knee health"

because it reflects aspects of pain, symptom severity, and ADL including movements or activities that are directly relevant and difficult for patients with advanced knee OA. When tested for construct validity against KOOS and WOMAC domains, the KOOS, JR compared favorably. We are currently in the process of conducting a prospective validation study which will evaluate the use of the KOOS, JR in the short-term and mid-term.

Responsiveness

The high responsiveness of the KOOS, JR was not surprising considering that the seven questions represent activities or movements that are directly relevant to patients with knee OA and are difficult for these patients to perform before TKA. The KOOS, JR, however was derived from an instrument developed for patients with less-severe knee disability. Fewer subjects are needed to adequately power outcomes studies using highly responsive instruments. The possibility of reaching statistical significance with a smaller cohort may be an unintended benefit for future studies performed using the KOOS, JR. The confidence intervals around the KOOS, JR scores also were narrower than those of other domains, likely owing to elimination of all gender-based or culturally dependent items, which also may reduce sample size needs for comparative research projects.

Comparison With the KOOS-PS

Perruccio et al. [17] reported on the development and validation of the KOOS-PS, a short-form survey derived from the KOOS, also using Rasch analysis. However, that item was developed as a short-form physical function PROM, excluding KOOS pain domain items. Given that pain is an overwhelming consideration for patients considering TKA, the KOOS-PS is less responsive and less relevant to patients undergoing TKA. Although they also arrived at a seven-item survey, only three items overlapped with ours: rising from sitting, bending to floor, and twisting/pivoting on your knee [17]. Because our cohort included only patients undergoing primary TKA, the KOOS, JR is a survey instrument that is specifically focused on end-stage OA and its treatment. The KOOS, JR also was more responsive than the KOOS-PS while having a high correlation with that instrument, albeit with a higher ceiling effect owing to the lower demand of KOOS, JR activities.

The KOOS, JR is an efficient alternative to traditional knee PROMs, appropriate for clinical outcomes assessment or as a research tool. Some research projects, however, may

require the full KOOS or other long-form PROMs. However, only seven questions to complete may help make data capture more efficient and better suited for the frenetic clinical environment in the busy specialist's practice and help fill the increasing demand for comparative outcomes data. As the CMS and other payers continue moving toward tying pay-for-performance outcomes, the KOOS, JR and other such efficient, accurate PROMS will become increasingly important in assuring compliance with coming regulations.

References

1. Anderson JC, Gerbing DW. Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. *J Appl Psychol*. 1991;76:732–740.

2. Ayers DC, Zheng H, Franklin PD. Integrating patient-reported outcomes into orthopaedic clinical practice: proof of concept from FORCE-TJR. *Clin Orthop Relat Res*. 2013;471:3419–3425.

3. Bellamy N. WOMAC: a 20-year experiential review of a patient-centered self-reported health status questionnaire. *J Rheumatol*. 2002;29:2473–2476.

4. Bond M, Davis A, Lohmander S, Hawker G. Responsiveness of the OARSI-OMERACT osteoarthritis pain and function measures. *Osteoarthritis Cartilage*. 2012;20:541–547.

5. CMS.org. Centers for Medicare & Medicaid Services. Measure Methodology. Available at: http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Measure-Methodology.html. Accessed January 11, 2014.

6. Conaghan PG, Emerton M, Tennant A. Internal construct validity of the Oxford Knee Scale: evidence from Rasch measurement. *Arthritis Rheum*. 2007;57:1363–1367.

7. Fisher WP Jr. Reliability, separation, strata statistics. *Rasch Measurement Transactions*. 1992;6:238.

8. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess*. 1998;2:i-iv, 1–74.

9. Franklin PD, Harrold L, Ayers DC. Incorporating patient-reported outcomes in total joint arthroplasty registries: challenges and opportunities. *Clin Orthop Relat Res*. 2013;471:3482–3488.

10. Higgins LD, Taylor MK, Park D, Ghodadra N, Marchant M, Pietrobon R, Cook C; International Knee Documentation Committee. Reliability and validity of the International Knee Documentation Committee (IKDC) Subjective Knee Form. *Joint Bone Spine*. 2007;74:594–599.

11. Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *J Bone Joint Surg Am*. 2007;89:780–785.

12. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care*. 1980;28:632–642.

13. Linacre JM. A User's Guide to Winsteps-Ministep: Rasch-model Computer Programs. Program Manual 3.68.0. Available at: http://www.winsteps.com/manuals.htm. Accessed December 1, 2015.

14. Lyman S, Lee YY, Franklin PD, Li W, Mayman DJ, Padgett DE. Validation of HOOS, JR: a short-form hip replacement survey. *Clin Orthop Relat Res*. 2016. doi:10.1007/s11999-016-4718-2

15. Masters GN. A Rasch model for partial credit scoring. *Psychometrika*. 1982;47:149–174.

16. McCarthy CJ, Oldham JA. The reliability, validity and responsiveness of an aggregated locomotor function (ALF) score in patients with osteoarthritis of the knee. *Rheumatology (Oxford)*. 2004;43:514–517.

17. Perruccio AV, Stefan Lohmander L, Canizares M, Tennant A, Hawker GA, Conaghan PG, Roos EM, Jordan JM, Maillefert JF, Dougados M, Davis AM. The development of a short measure of physical function for knee OA KOOS-Physical Function Shortform (KOOS-PS): an OARSI/OMERACT initiative. *Osteoarthritis Cartilage*. 2008;16:542–550.

18. Roos EM, Brandsson S, Karlsson J. Validation of the foot and ankle outcome score for ankle ligament reconstruction. *Foot Ankle Int*. 2001;22:788–794.

19. 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.

20. Rothman M, Burke L, Erickson P, Leidy NK, Patrick DL, Petrie CD. Use of existing patient-reported outcome (PRO) instruments and their modification: The ISPOR good research practices for evaluating and documenting content validity for the use of existing instruments and their modification PRO task force report. *Value Health*. 2009;12:1075–1083.

21. Smith AB, Rush R, Fallowfield LJ, Velikova G, Sharpe M. Rasch fit statistics and sample size considerations for polytomous data. *BMC Med Res Methodol*. 2008;8:33.

22. Steiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and Use*. London, UK: Oxford University Press; 2003.

23. Wainer H, Kiely GL. Item clusters and computerized adaptive testing: a case for testlets. *J Educ Meas*. 1987;24:185–202.

24. Wechsler S. *Statistics at Square One*. 9th ed. London, UK: BMJ Publishing Group; 1996.