

## Validation of the Chinese (Mandarin) Version of the Oxford Knee Score in Patients with Knee Osteoarthritis

Kai Lin BS, Liangxiao Bao BS, Jian Wang MD, Kimie Fujita PhD, Kiyoko Makimoto PhD, Xiaoyan Liao PhD

Received: 2 May 2017 / Accepted: 31 August 2017 / Published online: 7 September 2017  
© The Association of Bone and Joint Surgeons® 2017

### Abstract

**Background** With the increasing number of patients with knee osteoarthritis undergoing TKAs in China, there is a clear need for a valid, short, joint-specific patient-reported outcome measure such as the Oxford Knee Score (OKS). **Questions/purposes** To test the translated and cross-culturally adapted Chinese (Mandarin) version of the Oxford Knee Score (OKS-CV) and its (1) reliability, (2) construct validity, (3) dimensionality, and (4) responsiveness.

Each author certifies that he or she, or a member of his or her immediate family, has no funding or commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.

Funding was received from the National Natural Science Foundation of China (81400868) (XL), the Science and Technology Planning Project of Guangdong Province (2014A020212201) (XL), the Science and Technology Planning Project of Guangdong Province (2014A020209034) (JW), and the Health Department of Guangdong Government (A2014409) (XL).

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.

Southern Medical University, Nanfang Hospital approved the human protocol for this investigation, and each author certifies that all investigations were conducted in conformity with ethical principles of research.

This work was performed at Nanfang Hospital, Southern Medical University, Guangzhou, China.

**Electronic supplementary material** The online version of this article (doi:10.1007/s11999-017-5495-2) contains supplementary material, which is available to authorized users.

K. Lin, X. Liao (✉)  
Department of Nursing, Nanfang Hospital, Southern Medical University, No. 1838 Guangzhou Avenue North, Guangzhou 510515, China  
e-mail: liaoxy@smu.edu.cn

**Methods** Three native Chinese bilingual translators (a professional English translator, an experienced orthopaedic surgeon, an advanced-practice nursing specialist) translated the English-language OKS into Mandarin Chinese. A consensus panel created a synthesis of those efforts, which then was back-translated by two bilingual nonmedical, professional English-language translators. The OKS-CV was developed according to the guidelines of copyright holders. Between March 2013 and March 2015, 253 patients underwent TKAs. Among them, 114 Mandarin-speaking patients with knee osteoarthritis underwent primary unilateral TKA (age,  $67 \pm 7$  years; range, 55–84 years; female, 80%; preoperatively 54% had moderate to severe knee osteoarthritis), completed the preoperative questionnaires, and were followed up, with a mean post-operative followup of 2.7 years (SD, 0.5 years). Eligibility criteria were (1) patients with knee osteoarthritis who were scheduled to have a primary unilateral TKA, (2) patients who were fluent in Mandarin, and (3) consent to participate. The exclusion criteria were: (1) lack of understanding of Mandarin, and (2) inability to comprehend the questionnaires owing to cognitive impairment. To evaluate test-retest reliability, another group of 35 Mandarin-speaking outpatients with knee osteoarthritis (age,  $61 \pm 10$  years; range, 44–84 years; female, 77%) was recruited to complete the OKS-CV twice at a 1-week interval. Reliability was tested using Cronbach's alpha and intraclass correlation coefficient (ICC). Construct validity was evaluated using Spearman's rank correlation coefficient to quantify the correlations between the OKS-CV and the WOMAC,

L. Bao, J. Wang  
Department of Arthroplasty, Nanfang Hospital, Southern Medical University, Guangzhou, China

Short Form-8 Health Survey (SF-8<sup>TM</sup>), and EuroQol Group 5-Dimension Self-Report Questionnaire (EQ-5D). Exploratory factor analysis was performed to clarify dimensionality. The eigenvalue indicates the importance of each factor obtained from factor analysis. Responsiveness was determined by standardized response mean (SRM) and effect size (ES) from preoperative and postoperative scores of the OKS-CV. Floor and ceiling effects also were analyzed.

**Results** The internal consistency (Cronbach's alpha = 0.89) and test-retest reliability (ICC = 0.93; 95% CI, 0.87–0.97) proved good. Convergent construct validity was supported by moderate to strong correlations between the OKS-CV and the WOMAC ( $r = -0.80$ ,  $p < 0.001$ ), the SF-8<sup>TM</sup> physical component summary ( $r = 0.65$ ,  $p < 0.001$ ), and the EQ-5D usual activities ( $r = -0.41$ ,  $p < 0.001$ ) and mobility ( $r = -0.35$ ,  $p < 0.001$ ). There also were correlations between the OKS-CV and the SF-8<sup>TM</sup> mental component summary ( $r = 0.58$ ,  $p < 0.001$ ) and the EQ-5D anxiety/depression ( $r = -0.35$ ,  $p < 0.001$ ). The factor analysis yielded three factors with eigenvalues greater than 1. Responsiveness was excellent (SRM = 1.52; ES = 1.52). No floor or ceiling effect was observed.

**Conclusions** The OKS-CV showed good acceptability and psychometric properties for the intended population. Future studies are needed to evaluate the mental state of patients with knee osteoarthritis.

**Clinical Relevance** The OKS-CV appears to be a reliable, valid, and responsive instrument for Chinese patients with knee osteoarthritis. Based on these results we believe the OKS-CV can be used as a valuable tool for the assessment of patient-reported outcomes in Chinese patients with knee osteoarthritis before and after TKA.

## Introduction

Knee osteoarthritis is the most common degenerative joint disease in older adult populations, often affecting the patient's daily activities and quality of life [28]. According

to the sixth national population census of the People's Republic of China in 2010, there were 177.65 million people older than 60 years in China, accounting for 13.26% of the total population [43]. The overall prevalence of symptomatic knee osteoarthritis in the population 45 years or older in China reached 8.1% [53]. One would expect a portion of this aging population with symptomatic knee osteoarthritis to consider total knee arthroplasty (TKA) in the future, and patient-reported outcomes will be an important measure of success.

Several scoring systems in the Chinese language are available for assessing knee disorders, some of which have been scientifically validated, such as the new Chinese Knee Society Scoring System [38]. There is increasing interest in using patient-reported outcome measures (PROMS) in clinical practice across the spectrum of osteoarthritis. The Oxford Knee Score (OKS) is a disease-specific, PROM that is used to evaluate joint pain and physical function before and after TKA in patients with knee osteoarthritis [14, 16, 30]. The OKS has been shown to be valid, responsive, and reliable for measuring patients' perceptions of their knee disorder [11]. It correlates with the achievement of patients' expectations and satisfaction after TKA [1, 7–9, 32] or nonsurgical management [25] with no observer bias. The OKS has been reported to be an independent predictor of knee ROM after TKA [40]. Although not designed for this purpose, the OKS is increasingly used to assess eligibility for primary [12] or revision [47] surgery. The OKS is short and does not require a clinical review to obtain objective data, in contrast to some other functional scoring systems. These factors have resulted in its widespread use in large cohort studies and comparative trials that assessed outcomes after TKA and the long-term durability of knee components [20, 41]. Although there is a Singaporean Chinese version of the OKS that has been validated in Chinese-speaking Singaporeans with knee osteoarthritis [60], the Singapore Chinese version of the OKS may not be culturally adaptable to patients with knee osteoarthritis in mainland China, considering the sociocultural differences in lifestyle, language, and family support structures between Singapore and China.

We therefore sought to translate and cross-culturally adapt the OKS for use in mainland China, and test it for (1) reliability, (2) construct validity, (3) dimensionality, and (4) responsiveness.

## Methods

The medical ethics committee of Nanfang Hospital approved this study (NFEC-201209-K9), and all participants provided written informed consent.

K. Fujita  
Department of Health Science, Faculty of Medical Science,  
Kyushu University, Fukuoka, Japan

K. Makimoto  
Department of Nursing, Konan Women's University, Hyogo,  
Japan

K. Makimoto  
Department of Nursing, Graduate School of Medicine, Osaka  
University, Osaka, Japan

## Translation and Cross-Cultural Adaptation

The cross-cultural adaptation of the OKS to Chinese was performed according to the ISIS Outcomes Translation and Linguistic Validation Process (Oxford University Innovation [formerly ISIS Innovation], Oxford, UK) and a published guideline [59]. Briefly, the original English version of the OKS was translated independently to Chinese by three native Chinese bilingual translators (a professional English translator [NL], an experienced orthopaedic surgeon [JW], and an advanced-practice nursing specialist [XL]). A consensus panel was formed to discuss the three preliminary translations, which resulted in the first synthetic version. The synthetic version then was back-translated by two bilingual nonmedical, professional English translators (ZW, DC), who were blinded to the original English version. This process was continued until a final version was produced that had no disagreements between the English and Chinese versions. A consensus meeting that included all members of the consensus panel and the translators involved in the original translation process reviewed all of the reports and reached consensus on any discrepancies, thereby finalizing the pretesting Chinese version.

After the translation process was complete, 20 Mandarin-speaking patients with knee osteoarthritis who were scheduled to undergo TKA were asked to (1) decide if all of the items on the questionnaire were understandable and (2) complete the questionnaire. These subjects comprised patients with different socioeconomic and educational backgrounds (mean age,  $63 \pm 6$  years; range, 52–75 years; male, 25%; living in rural area, 60%; unemployed, 50%; illiteracy, 20%; primary school, 35%). Each item and response option were read aloud to the patients and they were asked about their understanding of the meaning of the item and the chosen response. Responses were recorded by the interviewer. No difficulties encountered by the respondents were noted in the pretest phase. Therefore, no further modification was required for the pretesting version. The final Chinese version that had been approved at the consensus panel meeting and all the versions of the scale were sent to and approved by the copyright holders of the OKS (Oxford University Innovation Limited, Oxford, UK) (<https://process.innovation.ox.ac.uk/p/oks/stage/1/oxford-knee-score.aspx>).

## Study Design and Subjects

This is a prospective cohort study of patients scheduled to undergo TKA. Between March 2013 and March 2015, 253 patients underwent TKA at the department of arthroplasty of a tertiary referral hospital in Guangzhou,

China. We applied inclusion and exclusion criteria to this group, and excluded 139 patients (55%), leaving 114 (45%) who participated (Group 1) and whose surveys were analyzed for psychometric properties of the OKS-CV. A flow chart outlines the assignment of patient groups (Fig. 1) Eligibility criteria were (1) patients with knee osteoarthritis who were scheduled to have a primary unilateral TKA, (2) patients who were fluent in Mandarin, and (3) consent to participate. The exclusion criteria were: (1) lack of understanding of Mandarin, and (2) inability to comprehend the questionnaires owing to cognitive impairment. After consenting to the research, all patients were interviewed by a trained interviewer (KL) and asked to complete four questionnaires before their planned TKA: the OKS-CV, WOMAC [2], Short Form-8 Health Survey (SF-8™; Quality Metric Incorporated, Lincoln, RI, USA), and the EuroQol Group 5-Dimension Self-Report Questionnaire (EQ-5D) [3]. Sociodemographic data and the presence of chronic medical status also were recorded. With a mean postoperative followup of 2.7 years (SD, 0.5 years), 114 patients who completed the preoperative questionnaires were asked to complete the OKS-CV and the WOMAC again postoperatively.

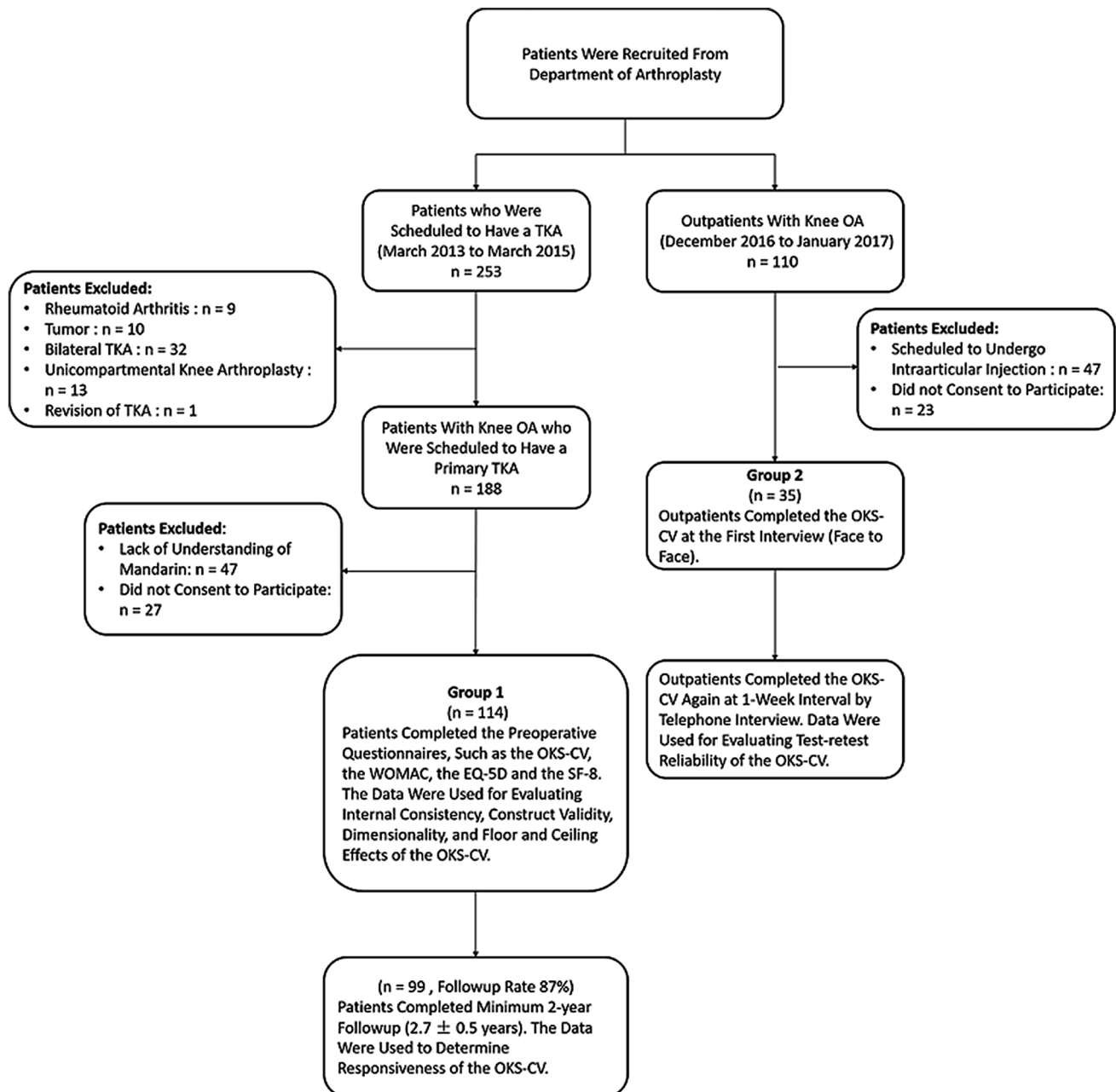
For assessing test-retest reliability of the OKS-CV, a different group of 35 outpatients with knee osteoarthritis (Group 2) was recruited between December 2016 and January 2017 at the same hospital. Eligibility and exclusion criteria for Group 2 were largely the same as those for Group 1. Because the measured condition of the patients was expected to change after TKA, “were scheduled to have a primary TKA” was removed from the eligibility criteria for Group 2. Moreover, to minimize changes of clinical status, patients who were scheduled to undergo intraarticular injection between the test-retest assessments were excluded. Among them, the OKS-CV was measured at the first interview (face to face) and again at a 1-week interval by telephone interview. The sample size for evaluating the test-retest reliability in this study is consistent with that of Naal et al. [42].

All patients were diagnosed with knee osteoarthritis by an orthopaedic surgeon (JW) according to radiographic and clinical criteria of the American College of Rheumatology diagnostic criteria [29]. There were no differences in the age, sex, and socioeconomic and educational backgrounds between the study sample and our routine patient population of the last 5 years ( $n = 715$ ) at our hospital (all  $p > 0.05$ ). As one of the three largest hospitals in Guangdong province, our hospital attracts patients from all cultures and income levels. Access to the hospital is open to every patient, and our patients are a mixture of urban and rural inhabitants. The study cohort therefore was considered representative.

## Patient Demographics and Clinical Background

A total of 114 patients underwent TKA (Group 1: mean age,  $67 \pm 7$  years; range, 55–84 years), completed the preoperative questionnaires, and were followed up by telephone interview, with a mean postoperative followup of 2.7 years (SD, 0.5 years). Among them, one patient died, 13 patients had moved and thus were lost from the sample, and one patient refused to answer the questionnaires because of hospitalization for pneumonia. The final cohort

consists of 99 patients (87% of the preoperative number of patients; mean age,  $67 \pm 7$  years; female, 80%). For assessing test-retest reliability, a different group of 35 outpatients (Group 2) completed the OKS-CV twice in 1 week. There were no differences in sex and socioeconomic or educational backgrounds between the groups ( $p > 0.05$ ). However, age ( $F = 17.3$ ,  $p < 0.001$ ), duration of osteoarthritis ( $F = 13.7$ ,  $p < 0.001$ ), and BMI ( $F = 24.3$ ,  $p < 0.001$ ) of Group 2 were smaller than those of Group 1 (Table 1). There were no missing data for any item of the



**Fig. 1** A flow chart shows the assignment of patient groups in our study. OKS-CV = Oxford Knee Score-Chinese Version; EQ-5D = EuroQol Group 5-Dimension Self-Report Questionnaire; SF-8 = Short-Form 8 Health Survey.

**Table 1.** Demographic characteristics of the subjects

Characteristics	Group 1 (n = 114)	Group 2 (n = 35)
Female, number (%)	91 (80)	27 (77)
Age (years), mean (SD)	67 (7)	61 (10) <sup>†</sup>
Duration of osteoarthritis (years), mean (SD)	9 (6)	6 (4)*
Years of education, number (%)		
≤ 6	63 (55)	23 (66)
7–12	36 (32)	8 (23)
≥ 13	15 (13)	4 (11)
Profession, number (%)		
Worker	6 (5)	5 (14)
Retired	61 (54)	14 (40)
Unemployed	47 (41)	16 (46)
BMI, mean (SD)	27 (3)	24 (3) <sup>†</sup>
Chronic comorbidity, number (%)	78 (68)	NA
Covered by insurance, number (%)	103 (90)	31(89)
Living arrangement, number (%)		
Live alone	9 (8)	NA
With family	105 (92)	NA
Living area, number (%)		
Rural	44 (39)	17 (49)
Urban	70 (61)	18 (51)
Housing, number (%)		
Building with elevator	41 (36)	NA
Building without elevator	42 (37)	NA
Bungalow with stairs	16 (14)	NA
Bungalow without stairs	15 (13)	NA
Type of toilet, number (%)		
Western	81 (71)	NA
Squat	33 (29)	NA

Categorical variables = number (%); chronic comorbidity: hypertension (n = 50), hyperlipidemia (n = 16), diabetes mellitus (n = 20), back pain (n = 8), periartthritis of shoulder (n = 2), cardiovascular (n = 10), or osteoporosis (n = 16) and other conditions (n = 8), including chronic nephrosis, anemia, cholelithiasis, gout, and cervical vertebra disease; \*p < 0.05; †p < 0.001.

OKS-CV. Patients did not report any difficulty understanding or completing the questionnaire.

### Measures Tested

The OKS is a joint-specific, patient-reported, outcome measures instrument containing 12 items regarding knee symptoms and function, each with five levels. According to the original scoring system, the score of each item ranges from 1 to 5, and a summary score ranges from 12 (least difficult) to 60 (most difficult) [14]. According to a revised

scoring system [41], the score of each item ranges from 0 (severe problems) to 4 (no problems), with a summary score ranging from 0 (worst outcome) to 48 (best outcome). We used the revised scoring system. Because it is difficult to compare original and revised scores among studies, we collected radiologic data of all 114 participants in Group 1 from their medical records to get a better sense of the degree of osteoarthritis and establish external validity with other patient populations. The Kellgren and Lawrence grading system was used to classify the severity of osteoarthritis [33] (Supplemental Table 1. Supplemental material is available with the online version of *CORR*®.)

The WOMAC is a 24-item, disease-specific, functional measurement commonly used to assess symptoms and physical disability in patients with knee and/or hip osteoarthritis [2]. The WOMAC consists of three domains. Several versions of the WOMAC can be used. We used the five-point Likert rating system (each item is scored 0 to 4): stiffness (two items; range, 0–8), pain (five items; range, 0–20), physical function (17 items; range, 0–68). The total score is calculated by summing the three domain scores (range, 0–96), with higher scores reflecting worse pain, stiffness, and physical function. The Chinese version of the WOMAC has been validated for use in China [51].

The EuroQol Group 5-Dimension Self-Report Questionnaire (EQ-5D) is a self-reporting questionnaire that measures quality of life using a self-classifier and a VAS (EQ-VAS) [46]. The self-classifier contains five dimensions—mobility, self-care, pain/discomfort, usual activities, and anxiety/depression—with three levels of response to each dimension. Utility valuations for all 243 EQ-5D health states are based on time tradeoff evaluations by 1222 members of the general public in China [39]. The EQ-VAS is a vertical, graduated (0–100 points) 20-cm VAS, where 100 represents the “best imaginable health state” and 0 the “worst imaginable health state” [61]. The Chinese version of EQ-5D has been validated for use in China [57].

The SF-8 is derived from the SF-36® for the purpose of yielding comparable scores for the eight health dimensions and two summary measures of the SF-36 with minimal respondent burden [58]. It consists of eight items: (1) general health, (2) physical functioning, (3) role physical, (4) bodily pain, (5) vitality, (6) social functioning, (7) emotional roles, and (8) mental health. Each item uses a five- or six-point Likert-type scale. All eight items are classified in two summaries, with the physical component summary generated by items 1 to 4 and the mental component summary produced from items 5 to 8 [36, 56]. The item scores and summary scores are calibrated on a scale of 0 to 100, with higher scores reflecting better subject status. The Chinese version of SF-8 has been validated for use in China.

### Reliability Testing of the OKS-CV

Internal consistency reflects the strength of the relationship among the items in the scale. In the current study, Cronbach's alpha was used to assess internal consistency of the OKS-CV. Internal consistency was deemed satisfactory if this index was 0.7 to 0.9 [30]. The coefficient also was calculated by eliminating one item from each of the 12 questions. All items were examined for correlation with the overall score [54]. The items should be at least moderately correlated with each other, and each item should correlate with the total scale score by greater than 0.30 [55]. The scores for the analysis of internal consistency of the OKS-CV were completed at administration of the pre-TKA assessment in Group 1 ( $n = 114$ ).

Test-retest reliability, referring to the reproducibility of the measurement, gives an indication of the stability of the test instrument with time. The test-retest reliability of the OKS-CV involved 35 outpatients in Group 2, and was assessed using an intraclass correlation coefficient (ICC) from a one-way random effects model. The ICC ranges from 0 (no agreement) to 1 (perfect agreement), with an ICC greater than 0.80 being considered an indicator of high reproducibility [34]. To ensure that the patients have a relatively stable clinical condition, the test-retest reliability of the OKS-CV was assessed using a recommended interval of 1 week between two measurements [18, 45, 52, 55].

### Construct Validity Testing of the OKS-CV

The construct validity was examined by means of convergent validity and divergent validity evaluations [35]. Construct validity refers to the degree to which the scale assesses the underlying theoretical construct it is supposed to measure. Convergent validity and divergent validity are two aspects of construct validity. To assess the construct validity of the OKS-CV, Spearman's rank correlation coefficients were calculated between the domains of the OKS-CV and WOMAC, and the related EQ-5D and SF-8 subscores. Hypotheses were formulated regarding the expected magnitude and direction of relations between the subscales of the OKS-CV and the above-mentioned instruments. For convergent validity, three theories were proposed. The OKS-CV should (1) have moderate to strong correlations with the pain and physical function domains of the WOMAC; (2) have moderate to strong correlations with the SF-8 physical component summary [14, 21]; and (3) have moderate to strong correlations with EQ-5D mobility and usual activities. For divergent validity, we postulated that (1) the correlation between the OKS-CV and EQ-5D VAS would be less than 0.35; and (2) the OKS-CV was expected to be least correlated with EQ-5D

anxiety/depression and the SF-8 mental component summary. All scores for the analysis of construct validity of the OKS-CV were completed at administration of the pre-TKA assessments in Group 1 ( $n = 114$ ).

### Dimensionality Testing of the OKS-CV

Dimensionality, referring to a measure of whether all items in a subscale relate to a single latent variable, was assessed using the exploratory factor analysis with nonorthogonal promax rotation [22, 60]. The exploratory factor analysis consists of two steps, namely finding an initial solution, and then rotating that solution [31]. Each retained factor had eigenvalues greater than 1 [19]. Eigenvalues indicate the importance of each factor in explaining the variability and correlations in the observed sample of data. Items with a factor loading of 0.40 or greater were considered acceptable [37]. Factor loadings represent how much a factor explains a variable in factor analysis. The scores for the analysis of dimensionality of the OKS-CV were completed at administration of the pre-TKA assessment in Group 1 ( $n = 114$ ).

### Responsiveness Testing of the OKS-CV

Responsiveness refers to a measurement property concerning change scores that might be used to assess improvement or deterioration on a health status. The responsiveness of the OKS-CV was evaluated by calculating standardized response mean (SRM) and effect size (ES) in patients who completed pre- and post-TKA assessments of the OKS-CV ( $n = 99$ , 87% of the patients pre-TKA in Group 1). The SRM is defined as the mean change scores between preintervention and postintervention divided by the SD of the change scores. The ES is defined as the observed change in scores between preintervention and postintervention divided by the SD of the baseline (preintervention) score. It has been suggested that ES of 0.2, 0.5, and 0.8 are regarded as small, medium, and large degrees of change, respectively [14]. A previous study suggested that function improvement may reach a plateau 1 year after TKA and is maintained at 3 years after TKA [50]; the followup of our study cohort (mean, 2.7 years; SD, 0.5 years) therefore was considered to be long enough after TKA that the patients would have been expected to reach maximal recovery.

We also examined the distribution of floor and ceiling effects of the OKS-CV in the patients who completed the pre-TKA assessment of the OKS-CV in Group 1 ( $n = 114$ ) by analyzing the proportion of individuals obtaining the lowest (0 points) and highest (48 points) scores,

respectively. The presence of floor or ceiling effect usually is defined as 15% of individuals in a sample achieving the extreme of the range (maximum or minimum level) [42]. The effects represent a measuring limitation, indicating that a scale may not be possible to measure a meaningful improvement or deterioration of a target condition.

### Sample Size Determination

The required sample size was determined based on the assumptions made for establishing the convergent validity. The convergent validity was established via Spearman's rank correlation coefficients, reflecting the strength of the association between the OKS-CV and the criterion-related scales. A moderate effect size of 0.30 for convergent validity was assumed [10], and the power and type 1 error probability set at 0.80 and 0.05 [6], respectively. With these assumptions, the required number of participants was 41 [6].

### Statistical Analyses

Demographic data were summarized using descriptive statistics, with the mean and SD obtained for continuous variables and the percentage for categorical variables. Internal consistency was estimated by using Chronbach's alpha. The test-retest reliability was assessed by using ICC. Convergent validity and divergent validity indicate the scale's construct validity. Spearman's rank correlations were used because of the nonparametric nature of the data. The correlation was considered strong, moderate, or weak if the coefficient was greater than 0.5, 0.35 to 0.5, or less than 0.35, respectively [15]. Higher correlation coefficients are expected for convergent validity, and lower correlation coefficients are expected for divergent validity [16, 23]. Dimensionality was assessed using exploratory factor analysis. The Kaiser-Meyer-Olkin test and Bartlett's test of sphericity were performed to confirm sample and items adequacy. The Kaiser-Meyer-Olkin measure should be greater than 0.5, and the significance level (*p*) for Bartlett's test should be less than 0.05 for a satisfactory factor analysis to proceed. Responsiveness was evaluated by using the SRM and ES, and then using paired *t* tests to compare the change in the scores from preoperative to postoperative. Correlations were calculated with bivariate two-tailed Pearson's correlations. Missing data were treated as described by Nilsdotter et al. [44]. The percentage of missing data was tabulated for each item; less than 5% was considered acceptable. All tests were two-tailed. A probability less than 0.05 was considered statistically significant. Data were analyzed with IBM SPSS statistics Version 19.0 software (IBM Corp, Armonk, NY, USA).

## Results

### Reliability

Internal consistency was good, with Cronbach's alpha at 0.885. Cronbach's alpha coefficients were slightly higher (0.888) after removing Item 8 "Pain in bed at night." The corrected item-total correlations exceeded 0.3 for all items (Table 2).

All 35 patients completed test-retest reliability. Mean scores for the first and second administrations of the OKS-CV were  $30 \pm 9$  and  $32 \pm 8$ , respectively. The OKS-CV showed excellent relative reliability with an ICC of 0.93 (95% CI, 0.87–0.97) for the total score.

### Construct Validity

Three a priori assumptions of convergent validity and one assumption of divergent validity were confirmed with moderate to strong correlations ( $r > 0.35$ ) between the OKS-CV and the WOMAC, the SF-8 physical component summary, the EQ-5D mobility and usual activities, and a weak correlation ( $r < 0.35$ ) between the OKS-CV and EQ-5D VAS (Table 3). With respect to convergent validity,

**Table 2.** Reliability of the Chinese version of the Oxford Knee Score

Item	Mean $\pm$ SD	Item-total correlation	<i>p</i> Value	$\alpha$ if item removed
Usual level of knee pain	1.14 $\pm$ 0.89	0.57	< 0.001	0.87
Trouble with washing and drying	2.35 $\pm$ 0.97	0.60	< 0.001	0.87
Trouble with transport	2.01 $\pm$ 0.91	0.72	< 0.001	0.86
Walking duration before severe pain	1.65 $\pm$ 1.27	0.53	< 0.001	0.88
Pain on standing up from sitting	2.13 $\pm$ 1.00	0.69	< 0.001	0.86
Limping when walking	1.35 $\pm$ 1.24	0.57	< 0.001	0.87
Difficulty with kneeling	0.91 $\pm$ 1.18	0.49	< 0.001	0.88
Pain in bed at night	2.18 $\pm$ 1.27	0.30	0.001	0.89
Work interference attributable to pain	1.69 $\pm$ 1.07	0.80	< 0.001	0.86
Sense of knee instability	2.96 $\pm$ 1.04	0.43	< 0.001	0.88
Doing household shopping alone	2.30 $\pm$ 1.38	0.69	< 0.001	0.86
Trouble walking down stairs	2.34 $\pm$ 1.17	0.68	< 0.001	0.86

**Table 3.** WOMAC, EQ-5D, or SF-8™ of the study population and its correlation to the OKS-CV

Instrument	Mean ± SD	OKS-CV r (95% CI)	p Value
WOMAC	46 ± 19	-0.80 (-0.9 to -0.65)	< 0.001
Stiffness	4 ± 2	-0.59 (-0.74 to -0.41)	< 0.001
Pain	10 ± 4	-0.71 (-0.82 to -0.55)	< 0.001
Function	33 ± 14	-0.76 (-0.86 to -0.6)	< 0.001
EQ-5D	0.54 ± 0.19	0.48 (0.31–0.62)	< 0.001
Mobility	0.11 ± 0.05	-0.35 (-0.5 to -0.17)	< 0.001
Self-care	0.07 ± 0.06	-0.26 (-0.44 to -0.09)	< 0.01
Usual activities	0.08 ± 0.05	-0.41 (-0.55 to -0.23)	< 0.001
Pain/discomfort	0.10 ± 0.05	-0.27 (-0.44 to -0.06)	< 0.01
Anxiety/ depression	0.06 ± 0.05	-0.35 (-0.51 to -0.17)	< 0.001
EQ-5D VAS	70 ± 19	0.30 (0.11–0.47)	0.001
SF-8 PCS	36 ± 15	0.65 (0.5–0.76)	< 0.001
General health	36 ± 19	0.50 (0.32–0.63)	< 0.001
Physical functioning	32 ± 17	0.57 (0.42–0.67)	< 0.001
Role physical	39 ± 20	0.58 (0.42–0.7)	< 0.001
Bodily pain	37 ± 19	0.49 (0.3–0.64)	< 0.001
SF-8 MCS	49 ± 20	0.58 (0.41–0.72)	< 0.001
Vitality	46 ± 18	0.37 (0.16–0.55)	< 0.001
Social functioning	50 ± 26	0.59 (0.43–0.71)	< 0.001
Emotional roles	53 ± 25	0.40 (0.21–0.57)	< 0.001
Mental health	49 ± 25	0.46 (0.28–0.61)	< 0.001

OKS-CV = Chinese version of the Oxford Knee Score; EQ-5D = the EuroQol Group 5-Dimension Self-Report Questionnaire; SF-8 = Short Form-8™ Health Survey; PCS = Physical component summary; MCS = Mental component summary.

there were moderate to strong correlations between the OKS-CV and all subscales of the WOMAC ( $r = -0.80$ ,  $p < 0.001$ ), the SF-8 physical component summary ( $r = 0.65$ ,  $p < 0.001$ ), and the EQ-5D mobility ( $r = -0.35$ ,  $p < 0.001$ ) and usual activities ( $r = -0.41$ ,  $p < 0.001$ ). Divergent construct validity was satisfied only by the weak correlation between the OKS-CV and EQ-5D VAS ( $r = 0.30$ ,  $p < 0.001$ ). The OKS-CV correlated strongly with the SF-8 mental component summary ( $r = 0.58$ ,  $p < 0.001$ ) and moderately with EQ-5D anxiety/depression ( $r = -0.35$ ,  $p < 0.001$ ).

**Dimensionality**

The exploratory factor analysis with nonorthogonal promax rotation method obtained three factors. The retained factors had eigenvalues greater than 1, and explained 66% of the

**Table 4.** Exploratory factor analysis of the Chinese version of the Oxford Knee Score

Item	Factor 1	Factor 2	Factor 3
Usual level of knee pain	-0.20	0.95*	0.15
Trouble with washing and drying	0.35	0.43*	0.07
Trouble with transport	0.65*	0.29	-0.05
Walking time before severe pain	0.32	0.54*	-0.36
Pain on standing up from sitting	0.06	0.81*	0.16
Limping when walking	0.28	0.51*	-0.02
Difficulty with kneeling	0.68*	-0.02	-0.07
Pain in bed at night	-0.19	0.31	0.86*
Work interference attributable to pain	0.64*	0.29	0.12
Sense of knee instability	0.60*	-0.30	0.59*
Doing household shopping alone	0.91*	-0.05	-0.04
Trouble walking down stairs	0.80*	0.02	0.02
Initial eigenvalue	5.50	1.30	1.13
Variance accounted for (%)	45.79	10.84	9.39

\*The pattern matrix resulted from principal component factoring with promax rotation with the questions loading more than 0.4.

general variance (Table 4). The Kaiser-Meyer-Olkin measure of sampling adequacy (0.85) and Bartlett’s test of sphericity (chi-square = 666;  $p < 0.001$ ) were satisfied. Six items loaded on factor 1 (Items 3, 7, 9, 10, 11, 12), which was labeled “Physical function” and showed a strong correlation with the physical function subscale of the WOMAC ( $r = -0.73$ ,  $p < 0.001$ ). Factor 2, composed of five items (Items 1, 2, 4, 5, 6), was labeled “Pain” and showed a strong correlation with the pain subscale of the WOMAC ( $r = -0.66$ ,  $p < 0.001$ ). Factor 3 had only one item (Item 8), which was defined as “Sleeping problem” and showed a moderate correlation with the EQ-5D pain/discomfort ( $r = -0.35$ ,  $p < 0.001$ ) and anxiety/depression ( $r = -0.36$ ,  $p < 0.001$ ) and the SF-8 emotional roles ( $r = 0.36$ ,  $p < 0.001$ ) and mental health ( $r = 0.39$ ,  $p < 0.001$ ).

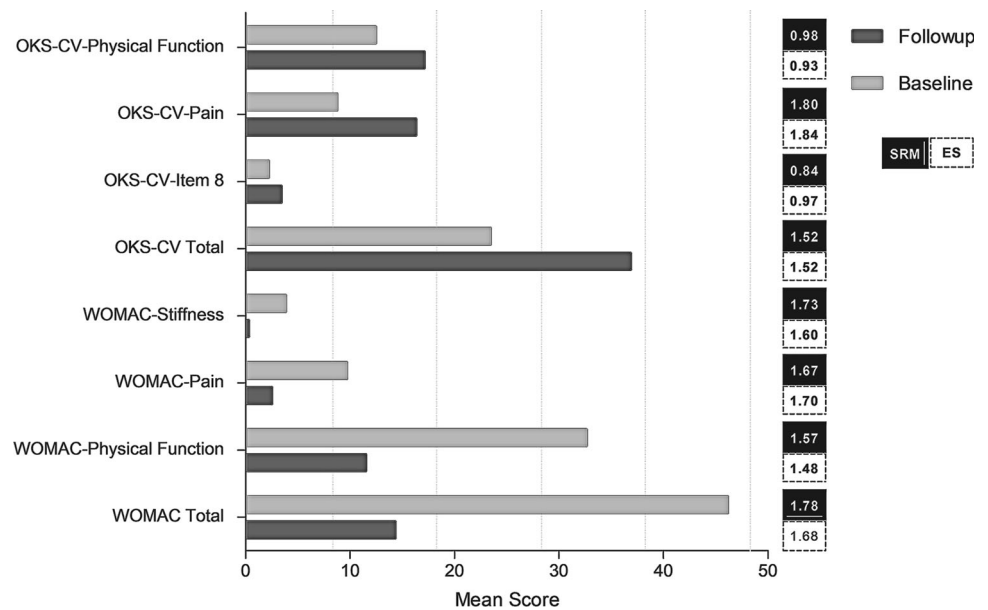
**Responsiveness**

The mean scores of the OKS-CV improved from 23 (SD, 9) to 37 (SD, 6) after surgery ( $t = -14.3$ ,  $p < 0.001$ ). The responsiveness values of the OKS-CV (SRM = 1.52, ES = 1.52) were close to those of the WOMAC (SRM = 1.78, ES = 1.68) (Fig. 2). The change scores of the OKS-CV were associated with the change scores of the WOMAC (Fig. 3).

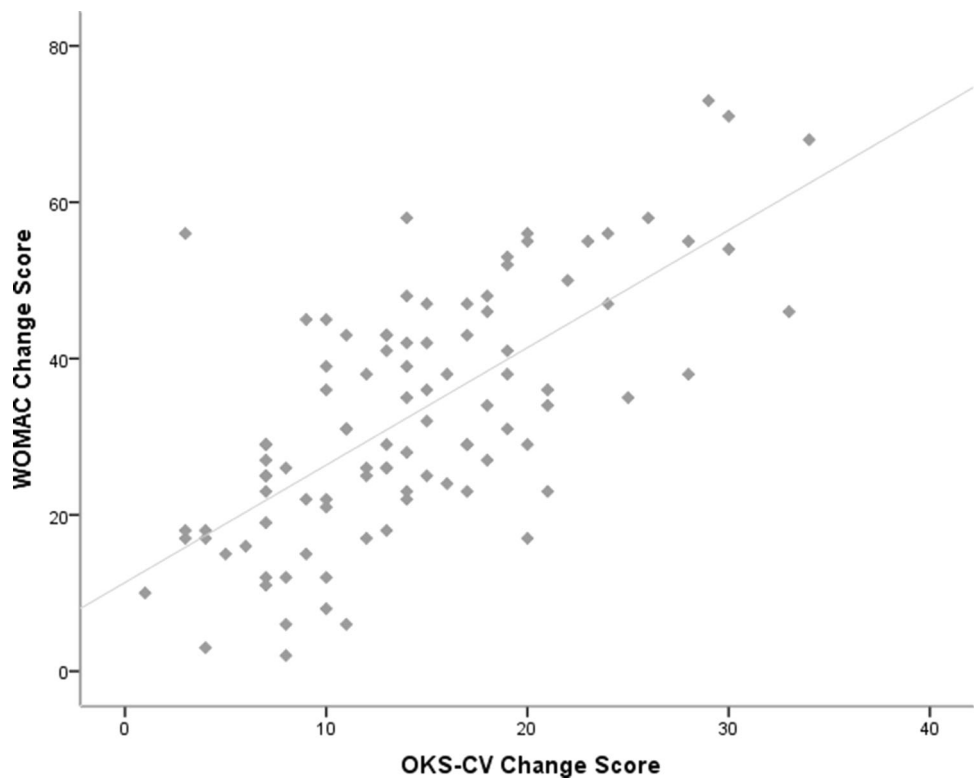
The results showed no floor or ceiling effect for the OKS-CV (Fig. 4). Two patients obtained the lowest score of 1 point (0.9%), and one patient obtained the highest score of 48 points (0.9%).



**Fig. 2** The standardized response means (SRM) and effect size (ES) measures at preoperative baseline and 2.7 (SD 0.5) years postoperative followup are shown (n = 99). OKS-CV = Oxford Knee Score-Chinese; Item 8 = Pain in bed at night.



**Fig. 3** The scatterplot shows the correlation between OKS-CV change score and WOMAC change score (r = 0.68, p < 0.001). A fitted linear regression line slopes upward from left to right. Change scores = postoperative score – preoperative scores; OKS-CV = Oxford Knee Score-Chinese Version.

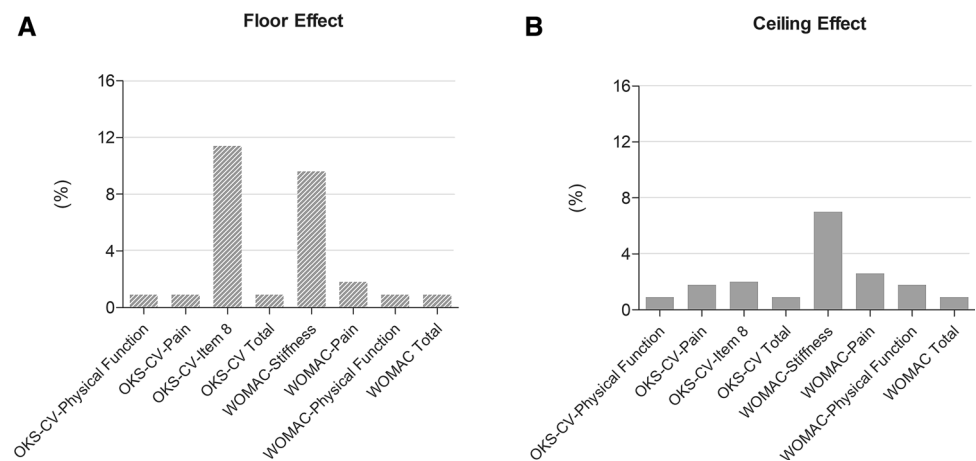


**Discussion**

From the cultural perspective, the Chinese mainland population has its own linguistic features, family structure, and lifestyle. For example, in our study, 92% patients lived with their adult offspring, 37% lived in a building without an elevator, 14% lived in a bungalow with stairs, and 29%

used squat toilets. This unique population is growing rapidly and the demand for TKA in an aging population can be expected to increase as well. PROMs are increasingly important tools to study the results of TKA in varied populations, and the OKS has been a favored instrument as it is easy for patients to complete, widely reported, and validated in many languages. Our study cross-culturally

**Fig. 4A–B** The graphs show the (A) floor and (B) ceiling effects in the OKS-CV and the WOMAC at preoperative baseline ( $n = 114$ ). OKS-CV = Oxford Knee Score-Chinese Version; Item 8 = Pain in bed at night.



adapted the OKS for use in mainland China, taking into account some of the unique aspects of the culture noted above, and validated it through standard psychometric pathways.

There are numerous limitations of the current study that should be acknowledged. First, our subjects were recruited from one arthroplasty center and represented mainly Chinese Mandarin-speaking patients, and this may limit generalizability elsewhere in mainland China. Second, we conducted a test-retest reliability assessment in a different group of patients who were younger and had a shorter duration of osteoarthritis than our main study cohort; our test-retest reliability result may not replicate in a larger representative sample. Third, the factor 3 (Item 8 ‘the pain at night’) obtained from the exploratory factor analysis may not be as robust as the other two factors, with the eigenvalue value close to 1. Fourth, the followup between the preoperative and postoperative scores used in our cohort was variable, since we used the longest followups possible for the OKS-CV measurements. A more-uniform followup would have provided more precise results. However, we considered this variability acceptable because we were judging the responsiveness of the OKS-CV and not specific eventual outcomes of TKA. In line with previous longitudinal studies suggesting that function improvement may reach plateaus 1 year after TKA and maintained at 3 years after TKA [50], there were no differences in the postoperative scores of the OKS-CV and the WOMAC during 2-year followup (24–35 months,  $n = 68$ ) and 3-year followup (36–46 months,  $n = 31$ ) in our cohort (Supplemental Table 2. Supplemental material is available with the online version of *CORR.*). The fifth limitation was that the patients were predominantly women and the sample size of this study restricts further subgroup analysis. Another larger prospective cohort study would be needed to address these issues.

Reliability, referring to the consistency or reproducibility of measurement, was evaluated by internal consistency and test-retest reliability. In published studies, Cronbach’s alpha coefficients of different language versions of the OKS range from 0.81 to 0.95 [5, 14, 16–18, 21, 26, 30, 42, 45, 48, 52, 55, 60], and ICCs range from 0.85 to 0.99 [16–18, 21, 26, 42, 45, 52, 55]. Regarding internal consistency, Cronbach’s alpha coefficient of the OKS-CV in this study was high (0.89), similar to the English version (0.87) [14] and cross-culturally adapted Italian, Thailand, French, Japanese, Portuguese, and Turkish versions [5, 21, 30, 45, 52, 55] (Supplemental Table 3. Supplemental material is available with the online version of *CORR.*). The high Cronbach’s alpha coefficient for the OKS-CV and acceptable corrected item-total coefficients for all 12 items confirmed that the OKS-CV is internally consistent, with the correspondent items properly correlated with each other. Regarding test-retest reliability, the value of the ICC (0.93; 95% CI, 0.87–0.97) indicated excellent test-retest reliability of the OKS-CV, similar to the Swedish, German, and Persian versions [16, 17, 42]. However, nearly 40% (six of 15) of published studies regarding translation and cultural adaptations of the OKS did not assess the test-retest reliability (Supplemental Table 3. Supplemental material is available with the online version of *CORR.*).

Evaluation of convergent validity and divergent validity was accomplished through testing of prespecified hypotheses regarding the relationship between the PROMS domain scores and evaluation systems already adequately validated. The convergent construct validity was good in this study. As hypothesized, the OKS-CV correlated moderately or strongly with domains in the WOMAC and SF-8 that measure similar health aspects. However, unlike the Singapore Chinese version [60], the correlations between the OKS-CV and the EQ-5D anxiety/depression

were much higher than expected, suggesting that knee osteoarthritis affects not only the physical but the mental health-related quality of life of our patients. Using a different scoring system, there were similar findings in the Swedish [16], and Singapore Chinese [60] versions. Potentially, pain and poor knee function may explain much of the emotional problems in patients with knee osteoarthritis. In support of this possibility, a study that developed a mapping algorithm between the OKS and EQ-5D utilities, based on a large sample size, suggested that the OKS predicted anxiety/depression responses quite accurately [13]. Furthermore, the potential reason that the OKS-CV was weakly correlated with the EQ-5D pain/discomfort in our study is that the patients were asked to report their status “during the past 4 weeks” in the OKS-CV, whereas in the EQ-5D it was to report their status as of “today”.

To our knowledge, there have been no consistent results among the published studies that have reported dimensionality. The items measuring pain and physical disability in the OKS did not load on the same factors [1, 18, 24]. In the current study, the results of exploratory factor analysis indicated a three-factor structure of the OKS-CV. First, factor 1 (composed of Items 3, 7, 9, 10, 11, 12 and labeled as “physical function”), which showed a strong correlation with the physical function subscale of the WOMAC, may indicate physical function. Second, factor 2 (composed of Items 1, 2, 4, 5, 6, and labeled as “pain”), which showed a strong correlation with the pain subscale of the WOMAC, may indicate pain. Item 8, “pain in bed at night,” loaded on the third factor independently. This finding was echoed by the moderate correlations between Item 8 and the EQ-5D pain/discomfort and anxiety/depression and the SF-8 emotional roles and mental health. This item may not perform as it was designed to. A possible explanation is the long average duration (years) of osteoarthritis in the patients in our study ( $9 \pm 6$  years), and the high percentage of patients (79%) choosing the response option “some nights,” “most nights,” or “every night”. Previous studies suggest that patients with osteoarthritis commonly report sleep disturbance that may be attributable to pain [4, 49]. The pain of osteoarthritis has negative effects on patients’ mood, sleep, and ability to participate in social and recreational activities [27]. Thus, our findings suggest that the OKS-CV may have a factor structure that encompasses more than two underlying health domains, with Item 8 representing a pain-related sleeping problem. Our findings also raise the possibility that it might be worthwhile to study the role of knee osteoarthritis in sleep and emotional problems, and the role of the OKS in predicting the anxiety/depression responses of the patients from a clinical perspective. Further studies with a larger sample size are needed to determine dimensionality of the scale in the

intended population, and to evaluate the mental state of patients with knee osteoarthritis.

In the current study, the corresponding values of the OKS-CV showed excellent responsiveness. Moreover, the change scores of the OKS-CV were associated with the change scores of the WOMAC. These results suggest that the OKS-CV is sensitive to measures of the patient’s pre-operative and postoperative changes, which can help clinicians and researchers determine whether clinical symptoms have improved after an intervention such as a TKA. However, only 13% (two of 15) of the published studies regarding translation and cultural adaptation of the OKS have conducted responsiveness testing (Supplemental Table 3. Supplemental material is available with the online version of *CORR*).

The Chinese version of the OKS is culturally and linguistically equivalent to the original English version. The results of this study support the use of the OKS-CV as a reliable, valid, responsive instrument for assessing PROMS outcomes in Chinese Mandarin-speaking patients with knee osteoarthritis. This version also will allow cross-cultural comparisons of PROMS outcomes among Chinese Mandarin-speaking patients with knee osteoarthritis to previously published studies from countries and populations in which the OKS was assessed in other languages.

**Acknowledgments** We thank Zhongqiang Wang PhD and Daiqiu Chen PhD (both from the English Department, College of Foreign Studies, Southern Medical University), and Ni Ling MD (Medical Center for Overseas Patients, Nanfang Hospital), for help in the translation process. In addition, we thank Qihong Li and Xiaodan Li for help with data collection, and the patients who participated in this study.

## References

1. Baker PN, van der Meulen JH, Lewsey J, Gregg PJ; National Joint Registry for England and Wales. The role of pain and function in determining patient satisfaction after total knee replacement: data from the National Joint Registry for England and Wales. *J Bone Joint Surg Br.* 2007;89:893–900.
2. 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.
3. Brooks R. EuroQol: the current state of play. *Health Policy.* 1996;37:53–72.
4. Campbell CM, Buenaver LF, Finan P, Bounds SC, Redding M, McCauley L, Robinson M, Edwards RR, Smith MT. Sleep, pain catastrophizing, and central sensitization in knee osteoarthritis patients with and without insomnia. *Arthritis Care Res (Hoboken).* 2015;67:1387–1396.
5. Charoencholvanich K, Pongcharoen B. Oxford knee score and SF-36: translation & reliability for use with total knee arthroscopy patients in Thailand. *J Med Assoc Thai.* 2005;88:1194–1202.

6. Chen J, Lu Q, Wu XY, An YZ, Zhan YC, Zhang HY. Reliability and validity of the Chinese version of the behavioral pain scale in intubated and non-intubated critically ill patients: two cross-sectional studies. *Int J Nurs Stud.* 2016;61:63–71.
7. Clement ND, Burnett R. Patient satisfaction after total knee arthroplasty is affected by their general physical well-being. *Knee Surg Sports Traumatol Arthrosc.* 2013;21:2638–2646.
8. Clement ND, MacDonald D, Patton JT, Burnett R. Post-operative Oxford knee score can be used to indicate whether patient expectations have been achieved after primary total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc.* 2015;23:1578–1590.
9. Clement ND, MacDonald D, Simpson AH. The minimal clinically important difference in the Oxford knee score and Short Form 12 score after total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc.* 2014;22:1933–1939.
10. Cohen J. A power primer. *Psychol Bull.* 1992;112:155–159.
11. Conaghan PG, Emerton M, Tennant A. Internal construct validity of the Oxford Knee Scale: evidence from Rasch measurement. *Arthritis Rheum.* 2007;57:1363–1367.
12. Dakin H, Gray A, Fitzpatrick R, MacLennan G, Murray D; KAT Trial Group. Rationing of total knee replacement: a cost-effectiveness analysis on a large trial data set. *BMJ Open.* 2012;2:e000332
13. Dakin H, Gray A, Murray D. Mapping analyses to estimate EQ-5D utilities and responses based on Oxford Knee Score. *Qual Life Res.* 2013;22:683–694.
14. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br.* 1998;80:63–69.
15. DeBette C, Parratte S, Maucourt-Boulch D, Blanc G, Pauly V, Lustig S, Servien E, Neyret P, Argenson JN. French adaptation of the new Knee Society Scoring System for total knee arthroplasty. *Orthop Traumatol Surg Res.* 2014;100:531–534.
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.
17. Ebrahimzadeh MH, Makhmalbaf H, Birjandinejad A, Soltani-Moghaddas SH. Cross-cultural adaptation and validation of the Persian version of the Oxford Knee Score in patients with knee osteoarthritis. *Iran J Med Sci.* 2014;39:529–535.
18. 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.
19. Fayers PM, Machin D. *Quality of Life: The Assessment, Analysis and Interpretation of Patient-reported Outcomes.* 2<sup>nd</sup> ed. Hoboken, NJ: John Wiley & Sons Inc; 2013.
20. Garratt AM, Brealey S, Gillespie WJ; DAMASK Trial Team. Patient-assessed health instruments for the knee: a structured review. *Rheumatology (Oxford).* 2004;43:1414–1423.
21. Goncalves RS, Tomas AM, Martins DI. Cross-cultural adaptation and validation of the Portuguese version of the Oxford Knee Score (OKS). *Knee.* 2012;19:344–347.
22. Guccione AA, Felson DT, Anderson JJ. Defining arthritis and measuring functional status in elders: methodological issues in the study of disease and physical disability. *Am J Public Health.* 1990;80:945–949.
23. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol.* 1993;46:1417–1432.
24. Harris K, Dawson J, Doll H, Field RE, Murray DW, Fitzpatrick R, Jenkinson C, Price AJ, Beard DJ. Can pain and function be distinguished in the Oxford Knee Score in a meaningful way? An exploratory and confirmatory factor analysis. *Qual Life Res.* 2013;22:2561–2568.
25. 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.
26. Haverkamp D, Breugem SJ, Sierevelt IN, Blankevoort L, van Dijk CN. Translation and validation of the Dutch version of the Oxford 12-item knee questionnaire for knee arthroplasty. *Acta Orthop.* 2005;76:347–352.
27. Hawker GA, Stewart L, French MR, Cibere J, Jordan JM, March L, Suarez-Almazor M, Goberman-Hill R. Understanding the pain experience in hip and knee osteoarthritis: an OARSI/OMERACT initiative. *Osteoarthritis Cartilage.* 2008;16:415–422.
28. Heidari B. Knee osteoarthritis prevalence, risk factors, pathogenesis and features: Part I. *Caspian J Intern Med.* 2011;2:205–212.
29. Hochberg MC, Altman RD, Brandt KD, Clark BM, Dieppe PA, Griffin MR, Moskowitz RW, Schnitzer TJ. Guidelines for the medical management of osteoarthritis: Part II: osteoarthritis of the knee. American College of Rheumatology. *Arthritis Rheum.* 1995;38:1541–1546.
30. Jenny JY, Diesinger Y. Validation of a French version of the Oxford knee questionnaire. *Orthop Traumatol Surg Res.* 2011;97:267–271.
31. Jolliffe IT, Morgan BJ. Principal component analysis and exploratory factor analysis. *Stat Methods Med Res.* 1992;1:69–95.
32. Judge A, Arden NK, Kiran A, Price A, Javaid MK, Beard D, Murray D, Field RE. Interpretation of patient-reported outcomes for hip and knee replacement surgery: identification of thresholds associated with satisfaction with surgery. *J Bone Joint Surg Br.* 2012;94:412–418.
33. Kellgren JH, Lawrence JS. Radiological assessment of osteoarthritis. *Ann Rheum Dis.* 1957;16:494–502.
34. Kim SJ, Basur MS, Park CK, Chong S, Kang YG, Kim MJ, Jeong JS, Kim TK. Crosscultural adaptation and validation of the Korean version of the New Knee Society Knee Scoring System. *Clin Orthop Relat Res.* 2017;475:1629–1639.
35. Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chronic Dis.* 1985;38:27–36.
36. Lefante JJ Jr, Harmon GN, Ashby KM, Barnard D, Webber LS. Use of the SF-8 to assess health-related quality of life for a chronically ill, low-income population participating in the Central Louisiana Medication Access Program (CMAP). *Qual Life Res.* 2005;14:665–673.
37. Lindeman RH, Merenda PF, Gold RZ. *Introduction to Bivariate and Multivariate Analysis.* Glenview, IL: Scott, Foresman; 1980.
38. Liu D, He X, Zheng W, Zhang Y, Li D, Wang W, Li J, Xu W. Translation and validation of the simplified Chinese new Knee Society Scoring System. *BMC Musculoskelet Disord.* 2015;16:391.
39. Liu GG, Wu H, Li M, Gao C, Luo N. Chinese time trade-off values for EQ-5D health states. *Value Health.* 2014;17:597–604.
40. Maempel JF, Clement ND, Brenkel IJ, Walmsley PJ. Range of movement correlates with the Oxford knee score after total knee replacement: a prediction model and validation. *Knee.* 2016;23:511–516.
41. Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr AJ, Dawson J. The use of the Oxford hip and knee scores. *J Bone Joint Surg Br.* 2007;89:1010–1014.
42. Naal FD, Impellizzeri FM, Sieverding M, Loibl M, von Knoch F, Mannion AF, Leunig M, Munzinger U. 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.
43. National Bureau of Statistics of China. The 2010 population census of the People's Republic of China. China Statistical

- Database (online). Available at: <http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/html/A0301a.htm>. Accessed July 8, 2013.
44. Nilsdotter AK, Lohmander LS, Klassbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS): validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord.* 2003;4:10.
  45. Padua R, Zanoli G, Ceccarelli E, Romanini E, Bondi R, Campi A. The Italian version of the Oxford 12-item Knee Questionnaire: cross-cultural adaptation and validation. *Int Orthop.* 2003;27:214–216.
  46. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med.* 2001;33:337–343.
  47. Reddy KI, Johnston LR, Wang W, Abboud RJ. Does the Oxford Knee Score complement, concur, or contradict the American Knee Society Score? *J Arthroplasty.* 2011;26:714–720.
  48. Reito A, Jarvisto A, Jansen E, Skytta E, Remes V, Huhtala H, Niemelainen M, Eskelinen A. Translation and validation of the 12-item Oxford knee score for use in Finland. *BMC Musculoskelet Disord.* 2017;18:74.
  49. Sasaki E, Tsuda E, Yamamoto Y, Maeda S, Inoue R, Chiba D, Okubo N, Takahashi I, Nakaji S, Ishibashi Y. Nocturnal knee pain increases with the severity of knee osteoarthritis, disturbing patient sleep quality. *Arthritis Care Res (Hoboken).* 2014;66:1027–1032.
  50. Sugita T, Kikuchi Y, Aizawa T, Sasaki A, Miyatake N, Maeda I. Quality of life after bilateral total knee arthroplasty determined by a 3-year longitudinal evaluation using the Japanese knee osteoarthritis measure. *J Orthop Sci.* 2015;20:137–142.
  51. Symonds T, Hughes B, Liao S, Ang Q, Bellamy N. Validation of the Chinese Western Ontario and McMaster Universities Osteoarthritis Index in patients from mainland China with osteoarthritis of the knee. *Arthritis Care Res (Hoboken).* 2015;67:1553–1560.
  52. Takeuchi R, Sawaguchi T, Nakamura N, Ishikawa H, Saito T, Goldhahn S. Cross-cultural adaptation and validation of the Oxford 12-item knee score in Japanese. *Arch Orthop Trauma Surg.* 2011;131:247–254.
  53. Tang X, Wang S, Zhan S, Niu J, Tao K, Zhang Y, Lin J. The prevalence of symptomatic knee osteoarthritis in China: results from the China Health and Retirement Longitudinal Study. *Arthritis Rheumatol.* 2016;68:648–653.
  54. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34–42.
  55. Tugay BU, Tugay N, Guney H, Kinikli GI, Yuksel I, Atilla B. Oxford Knee Score: cross-cultural adaptation and validation of the Turkish version in patients with osteoarthritis of the knee. *Acta Orthop Traumatol Turc.* 2016;50:198–206.
  56. Turner-Bowker DM, Bayliss MS, Ware JE Jr, Kosinski M. Usefulness of the SF-8 Health Survey for comparing the impact of migraine and other conditions. *Qual Life Res.* 2003;12:1003–1012.
  57. Wang H, Kindig DA, Mullahy J. Variation in Chinese population health related quality of life: results from a EuroQol study in Beijing, China. *Qual Life Res.* 2005;14:119–132.
  58. Ware JE, Kosinski M, Dewey JE, Gandek B. How to Score and Interpret Single-item Health Status Measures: A Manual for Users of the SF-8 Health Survey. Lincoln, RI: QualityMetric Incorporated; 2001.
  59. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, Erikson P; ISPOR Task Force for Translation and Cultural Adaptation. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health.* 2005;8:94–104.
  60. Xie F, Li SC, Lo NN, Yeo SJ, Yang KY, Yeo W, Chong HC, Fong KY, Thumboo J. 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.
  61. Zampelis V, Ornstein E, Franzen H, Atroshi I. A simple visual analog scale for pain is as responsive as the WOMAC, the SF-36, and the EQ-5D in measuring outcomes of revision hip arthroplasty. *Acta Orthop.* 2014;85:128–132.