



Translation and Validation of the Korean Version of the Forgotten Joint Score

Jangyun Lee, MD, Sang-Hyun Lim, MD, Du Hyun Ro, MD*, Myung Chul Lee, MD*, Hyuk-Soo Han, MD*

Department of Orthopedic Surgery, National Medical Center, Seoul,

**Department of Orthopedic Surgery, Seoul National University Hospital, Seoul, Korea*

Background: The Forgotten Joint Score (FJS) is a newly developed patient-reported outcome measure designed to evaluate clinical outcome after total knee arthroplasty (TKA). The FJS is known as a sensitive test with a low ceiling effect. It has been recently translated into many languages. However, no study has reported the validity or reliability of a Korean version of the FJS (K-FJS). Thus, the purpose of this study was to address this issue.

Methods: According to guidelines for cross-cultural adaptation, translation of the English version of the FJS was performed. After obtaining a license from the original developer, 150 patients who had undergone TKA at more than 1 year to less than 5 years ago completed the K-FJS, visual analog scale, Western Ontario and McMaster Universities Osteoarthritis index (WOMAC), and the 36-Item Short Form (SF-36) health survey. To measure test-retest reliability, the K-FJS was completed twice by telephone survey for 100 patients. Responsiveness was retrospectively calculated based on a survey of 50 patients at 3 months and 1 year after surgery.

Results: The K-FJS exhibited an excellent reliability (Cronbach's α , 0.967; intraclass correlation coefficient, 0.958; 95% confidence interval, 0.930–0.974). The ceiling effect of the K-FJS was 8.7% ($n = 13$), which was lower than the WOMAC's ceiling effect (10%). There was no floor effect. The correlation coefficients with WOMAC and SF-36 (physical function) were 0.708 and 0.682, respectively, indicating good construct validity. However, its correlation with mental health subscale of SF-36 was low ($r = 0.143$). At 3 to 12 months after TKA, the standardized response mean (SRM) was 0.67, which was lower than the SRM of WOMAC (1.03) obtained in the same period. The K-FJS demonstrated strong measurement properties in terms of good construct validity and reliability.

Conclusions: This study suggests that the K-FJS is an excellent instrument that can be used to monitor clinical outcomes after TKA. Using this standardized K-FJS, it would be possible for medical institutions to share more accurate clinical results.

Keywords: Knee, Osteoarthritis, Total knee arthroplasty, Forgotten joint score, Patient reported outcome measure

Total knee arthroplasty (TKA) is widely accepted as the standard therapy for end-stage degenerative osteoarthritis as it can provide pain relief and functional improvements.¹⁾ Postoperative outcome of TKA is often assessed based on surgeons' objective ratings rather than patients' subjective satisfaction. The gap between the two has led to the con-

cept of patient-reported outcome measures (PROMs).²⁾ Various tools such as health-related quality of life questionnaires have been developed to evaluate patient-centered clinical results.²⁾ Western Ontario and McMaster Universities Osteoarthritis index (WOMAC) was the first PROM developed in the early 1980s to assess patients with hip and knee osteoarthritis.³⁾ After rigorous validation, the WOMAC has been translated into more than 60 languages and used in many clinical studies related to the knee.³⁾ With improvement of surgical tools, implant materials, and techniques, average scores by the WOMAC and commonly used questionnaires are increasing. Many patients are now receiving the maximum scores on the WOMAC and Oxford knee score, indicating the presence of ceiling

Received August 31, 2020; Revised December 16, 2020;

Accepted December 16, 2020

Correspondence to: Hyuk-Soo Han, MD

Department of Orthopedic Surgery, Seoul National University Hospital,
101 Daehak-ro, Jongno-gu, Seoul 03080, Korea

Tel: +82-2-2072-4060, Fax: +82-2-764-2718

E-mail: oshawks7@snu.ac.kr

Copyright © 2021 by The Korean Orthopaedic Association

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Clinics in Orthopedic Surgery • pISSN 2005-291X eISSN 2005-4408

effects.⁴⁾ In an attempt to reduce ceiling effects, Behrend et al.⁵⁾ proposed a new disease-specific PROM known as the Forgotten Joint Score (FJS) in 2012. The FJS can be used to assess post-arthroplasty joint awareness, using 12 equally weighted questions that consider patients' ability to forget the artificial joint in everyday life as an ultimate outcome of arthroplasty. Recent studies have demonstrated that the FJS has higher reliability, better validity, and a lower ceiling effect than the WOMAC score does.⁵⁾ First developed in Switzerland, the FJS has been translated into many languages (including English, German, Japanese, and Danish) and successfully validated.⁶⁻¹⁰⁾ Recently, Adriani et al.¹¹⁾ have performed a systematic review focusing on the utility of the FJS and demonstrated that it has good construct validity and test-retest reliability. In Korea, many studies have individually analyzed the FJS score along with other PROMs. For instance, Kim et al.¹²⁾ have reported that FJS scores were higher in a unicompartmental knee arthroplasty group than in a TKA group. However, there is no consensus on the Korean version of the FJS (K-FJS). A validation study or a cross-cultural adaptation study of the FJS in Korea has not been reported. Creation and validation of a unified K-FJS hold a promise for its widespread application as a PROM tool for TKA patients. We believe that cross-cultural adaptation and conceptual equivalence are crucial to development of a K-FJS. Thus, the aim of this study was to develop a K-FJS that would be equally natural and acceptable as the original version. We further investigated the reliability, validity, and responsiveness of the K-FJS questionnaire to be used as a PROM tool in Korea.

METHODS

Translation and Cross-Cultural Adaptation

Translation and cross-cultural adaptation were proceeded in five steps according to guidelines of Guillemin et al.¹³⁾ and Wild et al.¹⁴⁾

Forward translation

The English version of the FJS was translated by two independent bilingual translators fluent in both English and Korean.

Synthesis of the translated versions

The two translated FJS questionnaires were merged into a single survey form. The merging process was focused on conceptual translation rather than literal translation. The final survey was revised to convey acceptable language for the broadest audience.

Backward translation

The questionnaire was translated back into English by two independent health professionals with English as their mother tongue. The two translators were familiar with terminologies of the area covered by this study.

Expert committee review including the licensor

A bilingual expert panel including the inventor of the questionnaire, original translators, and researchers of this study was convened to identify and resolve any inadequate expression or concept of the translation. The goal was to minimize any discrepancies between the final translation and the original questionnaire. Both the licensor and the licensee agreed to the revised version of the K-FJS.

Confirmation and pretesting

The pre-final version of the K-FJS was tested on 20 patients who underwent TKA. Respondent debriefing questions included what they thought about those questions and whether they could repeat those questions in their own words. The final version of the K-FJS is attached at the end of this manuscript (Appendix 1).

Validation Study

Reliability

Reliability refers to the degree to which a measurement tool is repeatable regardless of time and the tool can achieve consistent results. When measurement error is decreased, reliability is increased. We tested reliability by measuring whether the test was consistent across time (test-retest reliability) and across items (internal consistency). Internal consistency was measured using Cronbach's alpha as an index of whether items in one measurement tool were closely related to each other. Generally, a Cronbach's alpha of 0.80 and above indicates good internal consistency and a Cronbach's alpha of 0.90 and above indicates excellent internal consistency.¹⁵⁾ To measure test-retest reliability, a telephone questionnaire was conducted once again at 2 weeks to 1 month after the first survey. The time period was selected to be not too long so that the postoperative status was not changed. In the meantime, the time period was selected to be not too short so that patients could not recall the previous questionnaire. Intraclass correlation coefficient (ICC) was calculated to estimate test-retest reliability.¹⁵⁾

Validity

Validity refers to whether a measurement tool can accurately measure what the researcher intends to measure. Two general forms of validity, construct validity and con-

tent validity, were assessed for the K-FJS. Construct validity is the degree to which an instrument measures the trait or theoretical construct that it is intended to measure. To assess the construct validity, it was hypothesized that the K-FJS score would have a moderate to strong positive correlation with other PROM scoring systems (i.e., WOMAC and SF-36). Pearson's correlation coefficient was also calculated. Values greater than 0.6 indicated a strong correlation.¹⁶⁾ Content validity expresses how well the questionnaire covers all symptoms experienced by patients. The content validity was assessed for floor and ceiling effects. Floor effects include the proportion of patients scoring the lowest possible, whereas a ceiling effect expresses patients scoring the highest possible. Floor and ceiling effects of less than 15% of patients were considered to be adequate.¹⁷⁾

Responsiveness

Responsiveness measures sensitivity to changes within patients over time. High responsiveness means that the measurement tool is more sensitive in detecting changes within patients over time. Responsiveness to change was assessed using the standardized response mean (SRM) for the change between the 3-month follow-up time point and the 12-month follow-up time point. SRM was calculated as the average difference between two measurements divided by the standard deviation of differences between paired measurements, with higher SRM indicating greater responsiveness. According to the Cohen criteria, SRM of greater than 0.8, SRM of 0.5 to 0.8, and SRM of 0.2 or less indicate large, moderate, and small changes, respectively.¹⁸⁻²⁰⁾

Patient Selection

In accordance with the Institutional Review Board of Seoul National University Hospital (IRB No. H-1812-140-997), written informed consent was obtained from each patient prior to study. We retrospectively reviewed clinical databases at Seoul National University Hospital for degenerative osteoarthritis patients who had undergone primary TKA between January 2013 and December 2018. We identified 150 Korean-speaking patients whose average follow-up period was between 1 year and 5 years. We conducted a retrospective survey using the K-FJS, visual analog scale (VAS), WOMAC, SF-36, and Knee society score. For 100 patients, a telephone survey using the K-FJS was conducted once again at 3 weeks to 1 month after the first survey. In addition, among 150 patients, 50 patients with postoperative records of 3-month and 1-year follow-ups were compared. Score changes of VAS, WOMAC, and K-FJS were compared to investigate responsiveness over time. Patients with rheumatoid arthritis, posttraumatic os-

teoarthritis, history of previous knee surgery, and history of severe trauma or uncontrolled systemic disorders were excluded.

To minimize any difference in the level of understanding according to each patient's education level, a specialized orthopedic physician assistant (AJC) was assigned to help patients understand exact meanings of questionnaires before filling out. Both the interview-based survey and telephone survey were conducted by the same physician assistant to minimize error variance. Nonetheless, all questions were read by patients themselves. Intervention was minimized once the filling out process began. During this process, patients who could not read or understand Korean fluently were excluded from the analysis. General demographics of the patient population are summarized in Table 1.

Statistical Analysis

Study population was determined based on the standards proposed by Terwee et al.¹⁷⁾ A minimum of 100 patients were required for internal consistency analysis and a minimum of 50 patients were needed for analyzing floor or ceiling effects, reliability, and validity. All statistical analyses were conducted using IBM SPSS ver. 23.0 (IBM Corp., Armonk, NY, USA). Pearson's correlation coefficient for individual scoring system was calculated to study construct validity. Cronbach's alpha value, ICC, and SRM were used to determine internal consistency, test-retest reliability, and responsiveness, respectively, and 95% confidence intervals (CIs) were provided. Statistical significance was considered at p -value < 0.05.

Table 1. Demographic and Clinical Characteristics of Patients

Variable	Variable
Age (yr)	71.3 ± 6.5 (57–84)
Sex	
Female	124 (82.7)
Male	26 (17.3)
Side	
Right	74 (49.3)
Left	76 (50.7)
Time after surgery (mo)	20 ± 12 (12–60)

Values are presented as mean ± standard deviation (range) or number (%).

RESULTS

The K-FJS exhibited an excellent reliability (Cronbach's α , 0.967; ICC, 0.958; 95% CI, 0.930–0.974) (Table 2). The ceiling effect of the K-FJS was 8.7% ($n = 13$), which was lower than WOMAC's ceiling effect (10%). There was no floor effect (Table 3). Its correlation coefficients with WOMAC and SF-36 (physical function) were 0.708 and

0.682, respectively, indicating good construct validity (Figs. 1 and 2). At 3 to 12 months after TKA, the SRM of the K-FJS was 0.67, which was lower than the SRM of WOMAC (1.03) obtained in the same period (Table 4). Compared to WOMAC subset scores, the K-FJS had a high correlation with the pain subscale ($r = 0.561$) and the physical

Table 2. Measurement Properties of the K-FJS

Psychometric property	Value	<i>p</i> -value
Validity		
Construct validity [†]		
WOMAC score	0.708	< 0.001*
Knee Society Score		
Knee score	0.258	< 0.001*
Function score	0.889	< 0.001*
SF-36		
General health	0.149	0.074
Physical function	0.682	< 0.001*
Role physical	0.373	< 0.001*
Emotion physical	0.390	< 0.001*
Bodily pain	0.579	< 0.001*
Vitality	0.073	0.385
Social function	0.597	< 0.001*
Mental health	0.143	0.085
Content validity		
Ceiling effect	8.7 (13) [‡]	
Floor effect	-	
Reliability		
Internal consistency		
Cronbach's α	0.967	
Test-retest reliability		
Intraclass correlation coefficient	0.958 (0.930–0.974) [§]	< 0.001
Responsiveness		
Standardized response mean	0.67	

K-FJS: Korean version of the Forgotten Joint Score, WOMAC: Western Ontario and McMaster Universities Osteoarthritis index, SF-36: 36-Item Short Form.

*Statistically significant, $p < 0.05$. [†]Pearson's correlation coefficient. [‡]Percent (number). [§]95% Confidence interval.

Table 3. Ceiling and Floor Effects of the K-FJS and WOMAC

Variable	Ceiling effect	Floor effect
K-FJS	8.7 (13)	No
WOMAC	10 (15)	No

Values are presented as percent (number).

K-FJS: Korean version of the Forgotten Joint Score, WOMAC: Western Ontario and McMaster Universities Osteoarthritis index.

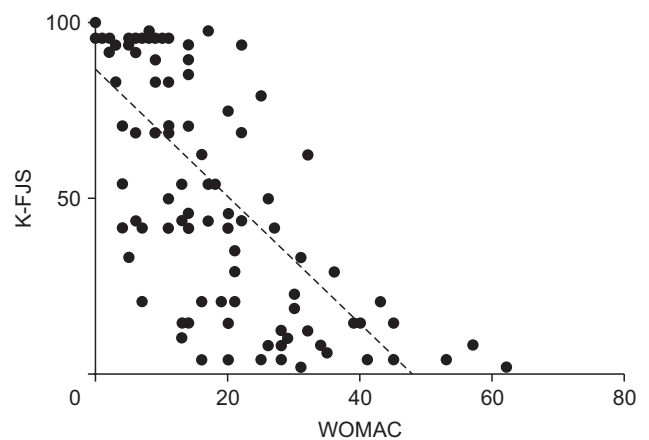


Fig. 1. Correlation of the Korean version of the Forgotten Joint Score (K-FJS) with the Western Ontario and McMaster Universities Osteoarthritis index (WOMAC).

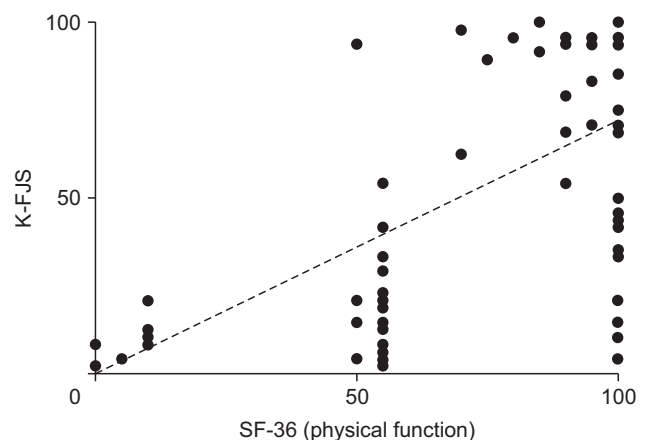


Fig. 2. Correlation of the Korean version of the Forgotten Joint Score (K-FJS) with the physical function subscale of the 36-Item Short Form (SF-36) health survey.

subscale ($r = 0.649$) and moderate correlation with the stiffness subscale ($r = 0.420$) (Table 5). However, its correlation with the mental health subscale of SF-36 was low ($r = 0.143$). Besides, albeit not a PROM, the Knee Society Function Score showed a strong correlation with the K-FJS ($r = 0.889$) (Fig. 3).

DISCUSSION

To minimize any difference in the level of understanding according to each patient's education level, researchers were assigned to help patients understand exact meanings of questionnaires before filling out. We believe this process is a good way to avoid half-hearted consistent answers or missing answers due to the lack of understanding. It also helps researchers to figure out where to improve during the early stage of the pilot study.

During the pilot study, many patients asked about the exact meaning of "Are you aware of your artificial joint?" A brief explanation was added to make sure that patients understood it as "How much are you aware of your artificial joint?" Afterwards, questions about confusing sentences were largely reduced. Also, as mentioned in the study of Cao et al.,²¹⁾ many patients mistakenly considered that items in the prefinal K-FJS were asking the frequency they were able to finish a corresponding activity

after TKA. Thus, we added the term "aware of" to each item in bold to minimize misunderstanding of questions.

According to a systematic review on the FJS using all articles that reported missing response percentages, question number 12 ("Are you aware of your artificial knee when doing your favorite sport?") had a significantly high missing response rate ($> 10\%$). Likewise, we found significantly high frequencies of responses saying that they had no favorite sport or it had been too long since they quitted exercise.^{7,9,21)} However, because we excluded all cases with any missing item, the exact missing response rate was not assessed. We believe this requires further investigation.

The validity study revealed a good correlation between the K-FJS and the WOMAC total score ($r = 0.708$, $p < 0.001$) (Fig. 1). This was similar to earlier cross-cultural studies reporting a high correlation between the FJS and the WOMAC score (Behrend et al.,⁵⁾ $r = 0.79$; Shadid et al.,⁹⁾ $r = 0.75$). Some studies have also shown less statistically significant results. The study conducted by Matsumoto et al.⁸⁾ showed a moderate correlation ($r = 0.52$) between Japanese version FJS and WOMAC score.

Table 4. Responsiveness of the VAS, K-FJS, and WOMAC Scores

Variable	Mean of change	SD	SRM*
VAS (0–10)	0.70	2.54	0.28
K-FJS	-10.64	15.98	0.67
WOMAC score	11.44	11.11	1.03

VAS: visual analog scale, K-FJS: Korean version of the Forgotten Joint Score, WOMAC: Western Ontario and McMaster Universities Osteoarthritis index, SD: standard deviation, SRM: standardized response mean.

*Mean postoperative score – mean preoperative score/standard deviation of the change in score.

Table 5. Pearson's Correlation Coefficient between the K-FJS and WOMAC Subscales

Variable	Ceiling effect	Floor effect	Correlation coefficient (r)
WOMAC (total)	10 (15)	No	0.708*
WOMAC pain	33 (50)	No	0.561*
WOMAC stiffness	33 (50)	No	0.420*
WOMAC physical function	14 (21)	No	0.649*

Values are presented as percent (number).

K-FJS: Korean version of the Forgotten Joint Score, WOMAC: Western Ontario and McMaster Universities Osteoarthritis index.

* $p < 0.001$.

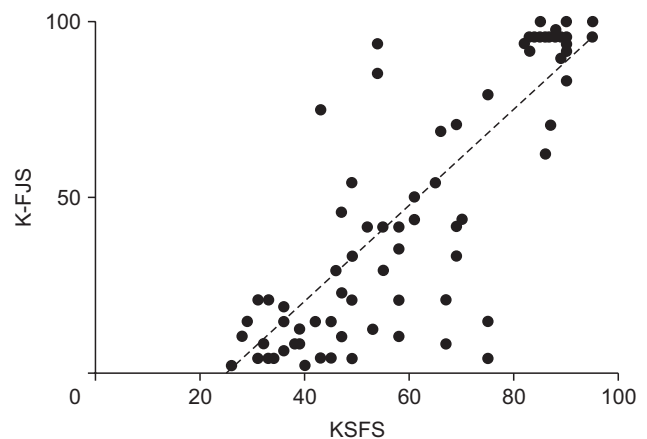


Fig. 3. Correlation of the Korean version of the Forgotten Joint Score (K-FJS) with the Knee Society Function Score (KSFS).

Compared to WOMAC subset scores, the K-FJS had a relatively weak correlation with the stiffness subscale ($r = 0.420$) (Table 5). Similar results have been previously reported in a Japanese study by Matsumoto et al.⁸⁾ ($r = 0.4$) and the original version study by Thomsen et al.¹⁰⁾ ($r = 0.52$). The WOMAC is composed of 24 items with three subscales (17 for physical function, 5 for pain, and 2 for stiffness). Two stiffness subscale questions include “stiffness after waking up in the morning” and “stiffness after sitting/lying or resting during the day.” However, there was no question in the original version FJS that specifically evaluated stiffness. The difference in questionnaire contents might have resulted in this weak correlation. Moreover, of a total of 96 points, the stiffness subscale was assigned a maximum of 8 points. Due to this small range, we believe that the stiffness subscale would not have much influence on the correlation between the K-FJS and the total WOMAC. Besides, the K-FJS had moderate correlations with SF-36 subscales of pain, physical function, and social function. This is comparable to the validation result of a Chinese version FJS presented by Cao et al.²¹⁾ Concerning these results and its correlations with WOMAC pain and WOMAC physical function subscales, the K-FJS showed a good convergent validity. In addition, the K-FJS showed a low correlation with SF-36 mental health subscale ($r = 0.143$). This reflects good discriminant validity, as highlighted in the English and Chinese versions of the FJS ($r = 0.23$ and $r = 0.086$, respectively).

No floor effect was observed in the total score of the K-FJS and WOMAC. We found a ceiling effect of 8.7% for the K-FJS as compared to 10% for the WOMAC (Table 3). Because the WOMAC questionnaire is composed of twice as many items as the FJS, it is harder for the WOMAC to have a ceiling effect. Considering that the K-FJS has less question numbers, we can conclude that the K-FJS has a lower ceiling effect. Earlier studies have presented that floor and ceiling effects lower than 15% are required for a study to have a reasonable content validity and that effects lower than 10% are considered ideal.¹⁷⁾ The ceiling effect of the K-FJS was 8.7%, meaning an ideal content validity. This result is better than findings presented in the first FJS study by Behrend et al.⁵⁾ (ceiling effect for the FJS and the WOMAC were 9.2% and 16.7%, respectively).

This study showed an excellent test-retest reliability (ICC, 0.958). High reliability is crucial to demonstrate the stability of a questionnaire over time. However, most other FJS validation studies have shown an ICC score range of 0.9 to 0.92.¹¹⁾ Unlike other studies, our test-retest was conducted over the telephone to relieve patients' discomfort and burdens. This procedure might have affected the ICC.

However, previous studies have reported that there is no statistical difference between in-person and telephone test-retest reliability.^{22,23)}

Internal consistency was measured using Cronbach's alpha value. The K-FJS received an excellent value of 0.967, which was comparable to the original version's value of 0.95.⁵⁾ High scores of internal consistency might mean homogeneity and reliability, indicating that items in the questionnaire were closely related to each other. However, too high Cronbach's alpha value (over 0.95) might reflect excessive duplication of contents. In such cases, researchers should be cautious when interpreting results.²⁴⁾ In the present study, we eliminated each item and re-evaluated Cronbach's alpha. The calculated Cronbach's alpha values ranged from 0.962 to 0.969. Internal consistency was not increased after removing certain items, and therefore, we could not find redundant items. According to a systematic review published by Adriani et al.,¹¹⁾ the mean Cronbach's alpha value of 10 reviewed articles was 0.95 (range, 0.91–0.98). Of these 10 articles, those published in English-speaking culture had a greater mean value of 0.97 (range, 0.95–0.98). Therefore, we can conclude that regardless of cross-cultural adaptation, high internal consistency is a characteristic of the FJS.

The responsiveness evaluated at 3 months and 1 year after surgery showed a moderate SRM (0.67). Moderate and high SRMs indicate sufficient internal responsiveness. Since our study design did not include preoperative K-FJS, we compared the change over time from 3 months to 1 year after surgery. Many patients did not show up at postoperative 6-month follow-up. According to a study by Hamilton et al.,⁷⁾ the effect size (Cohen's *d*) that compared change from 6 months to 12 months after surgery was 0.12. However, the effect size that compared change from preoperative data to 6 months after surgery was 2.6.⁷⁾ Our study results showed that the K-FJS was relatively sensitive in discriminating changes in clinical outcomes between 3 months and 1 year after surgery. We believe it is a suitable tool for monitoring clinical outcomes after surgery.

However, because clinical results might vary depending on how the follow-up period is selected, it is meaningful to compare the absolute SRM value with that of VAS or the WOMAC test result. In this study, the SRM of the K-FJS was higher than that of VAS (0.28), but lower than that of the WOMAC (1.03). This result is in accordance with a previous study by Bellamy et al.,³⁾ showing that the WOMAC score tends to show a higher responsiveness than other evaluation tools.

This study has several limitations. First, since preoperative FJS was not evaluated, we could not identify the

floor effect for the K-FJS. According to the study conducted by Hamilton et al.,⁷⁾ the floor effect of preoperative FJS in patients with TKA was 15%. Unlike Oxford knee score and WOMAC score, which initially targeted osteoarthritis patients, FJS-12 was developed to assess postoperative population. Such difference would have resulted in higher floor effect for the FJS.

Second, the mode of questionnaire administration was different for the two surveys. All the surveys were performed by a single orthopedic physician assistant who fully understood this study. However, the first survey was performed based on interviewer-administered mode and the second by telephone survey. According to previous research, interview format of survey or telephone survey is more effective for the elderly or patients with low social economic status.²⁵⁾ However, according to Lyons et al.,²⁶⁾ the interview format could systematically exaggerate health status compared with self-assessment. Likewise, patients may respond differently to K-FJS measures depending on the mode of questionnaire administration. Therefore, our study has a major limitation of not being able to confirm how the questionnaire results differ depending on the presence or absence of a researcher helping the questionnaire.

The K-FJS demonstrated strong measurement properties in terms of good construct validity and reliability. Our results suggest that it is an excellent instrument that can be used to monitor clinical outcomes after TKA. Using this standardized version of K-FJS, it would be possible for institutions to share more accurate clinical results.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGEMENTS

The authors thank Ae-Jin Choi (Seoul National University Hospital) for her contribution to this study.

ORCID

Jangyun Lee <https://orcid.org/0000-0001-5378-7001>
 Sang-Hyun Lim <https://orcid.org/0000-0001-5004-9775>
 Du Hyun Ro <https://orcid.org/0000-0001-6199-908X>
 Myung Chul Lee <https://orcid.org/0000-0002-8150-1573>
 Hyuk-Soo Han <https://orcid.org/0000-0003-1229-8863>

REFERENCES

- Hawker G, Wright J, Coyte P, et al. Health-related quality of life after knee replacement. *J Bone Joint Surg Am.* 1998;80(2):163-73.
- Kon E, Altadonna G, Filardo G, Di Matteo B, Marcacci M. Knee scoring systems. In: George Bentley, ed. *European surgical orthopaedics and traumatology: the EFORT textbook.* Heidelberg: Springer-Verlag Berlin Heidelberg; 2014. 3371-88.
- 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(12):1833-40.
- Marx RG, Jones EC, Atwan NC, Closkey RF, Salvati EA, Sculco TP. Measuring improvement following total hip and knee arthroplasty using patient-based measures of outcome. *J Bone Joint Surg Am.* 2005;87(9):1999-2005.
- Behrend H, Giesinger K, Giesinger JM, Kuster MS. The "forgotten joint" as the ultimate goal in joint arthroplasty: validation of a new patient-reported outcome measure. *J Arthroplasty.* 2012;27(3):430-6.e1.
- Baumann F, Ernstberger T, Loibl M, Zeman F, Nerlich M, Tibesku C. Validation of the German Forgotten Joint Score (G-FJS) according to the COSMIN checklist: does a reduction in joint awareness indicate clinical improvement after arthroplasty of the knee? *Arch Orthop Trauma Surg.* 2016;136(2):257-64.
- Hamilton DF, Loth FL, Giesinger JM, et al. Validation of the English language Forgotten Joint Score-12 as an outcome measure for total hip and knee arthroplasty in a British population. *Bone Joint J.* 2017;99(2):218-24.
- Matsumoto M, Baba T, Homma Y, et al. Validation study of the Forgotten Joint Score-12 as a universal patient-reported outcome measure. *Eur J Orthop Surg Traumatol.* 2015;25(7):1141-5.
- Shadid MB, Vinken NS, Marting LN, Wolterbeek N. The Dutch version of the Forgotten Joint Score: test-retesting reliability and validation. *Acta Orthop Belg.* 2016;82(1):112-8.
- Thomsen MG, Latifi R, Kallemose T, Barfod KW, Husted H, Troelsen A. Good validity and reliability of the forgotten joint score in evaluating the outcome of total knee arthroplasty. *Acta Orthop.* 2016;87(3):280-5.

11. Adriani M, Malahias MA, Gu A, Kahlenberg CA, Ast MP, Sculco PK. Determining the validity, reliability, and utility of the forgotten joint score: a systematic review. *J Arthroplasty*. 2020;35(4):1137-44.
12. Kim MS, Koh IJ, Choi YJ, Lee JY, In Y. Differences in patient-reported outcomes between unicompartmental and total knee arthroplasties: a propensity score-matched analysis. *J Arthroplasty*. 2017;32(5):1453-9.
13. 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(12):1417-32.
14. Wild D, Grove A, Martin M, et al. 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(2):94-104.
15. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19(4):539-49.
16. Evans JD. *Straightforward statistics for the behavioral sciences*. Pacific Grove: Thomson Brooks/Cole Publishing; 1996. 109-39.
17. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42.
18. Kim JG, Ha JK, Lee JY, Seo SS, Choi CH, Lee MC. Translation and validation of the Korean version of the international knee documentation committee subjective knee form. *Knee Surg Relat Res*. 2013;25(3):106-11.
19. Cohen J. Set correlation and multivariate methods. In: Cohen J, ed. *Statistical power analysis for the behavioral sciences*. Hillsdale, MI: Lawrence Erlbaum Associates; 1988. 467-530.
20. Chinsongkram B, Chaikeeree N, Saengsirisuwan V, Horak FB, Boonsinsukh R. Responsiveness of the Balance Evaluation Systems Test (BESTest) in people with subacute stroke. *Phys Ther*. 2016;96(10):1638-47.
21. Cao S, Liu N, Han W, et al. Simplified Chinese version of the Forgotten Joint Score (FJS) for patients who underwent joint arthroplasty: cross-cultural adaptation and validation. *J Orthop Surg Res*. 2017;12(1):6.
22. Garcia M, Rohlfs I, Vila J, et al. Comparison between telephone and self-administration of short form health survey questionnaire (SF-36). *Gac Sanit*. 2005;19(6):433-9.
23. Braunsberger K, Wybenga H, Gates R. A comparison of reliability between telephone and web-based surveys. *J Bus Res*. 2007;60(7):758-64.
24. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ*. 2011;2:53-5.
25. Sullivan LM, Dukes KA, Harris L, Dittus RS, Greenfield S, Kaplan SH. A comparison of various methods of collecting self-reported health outcomes data among low-income and minority patients. *Med Care*. 1995;33(4 Suppl):AS183-94.
26. Lyons RA, Wareham K, Lucas M, Price D, Williams J, Hutchings HA. SF-36 scores vary by method of administration: implications for study design. *J Public Health Med*. 1999;21(1):41-5.

슬관절 질문지 (Forgotten Joint Score-12)

환자 : _____ 날짜 : ____ . ____ . ____

귀하는 일상생활에서 건강한 관절을 인지하지 않고 지내게 됩니다. 하지만 아주 작은 문제라도 생기게 되면 관절을 인지하게 됩니다. 이것은 귀하가 관절을 생각하거나 그것에 주의를 기울인다는 것을 의미합니다. 다음 질문은 **일상생활에서 문제가 있는 슬관절을 얼마나 자주 인지하는지**에 관한 것입니다.

각 질문에 가장 알맞은 답을 선택하십시오.

귀하는 다음 상황에서 귀하의 슬관절에 대해 인지하고 있습니까?	전혀 인지하고 있지 않다	거의 인지하고 있지 않다	드물게 인지하고 있다	종종 인지하고 있다	거의 인지하 고 있다
1. 밤에 침대에 누워있을 때	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. 한 시간 이상 의자에 앉아있을 때	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. 15분을 넘게 걸을 때	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. 샤워하거나 목욕할 때	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. 차를 타고 여행할 때	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. 계단을 오를 때	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. 울퉁불퉁한 길을 걸을 때	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. 낮게 앉은 자세에서 일어날 때	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. 오랫동안 서 있을 때	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. 집안일을 하거나 정원일을 할 때	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. 산책하거나 가벼운 등산을 할 때	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. 좋아하는 운동을 할 때	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

© Copyright 2014 Behrend H, Giesinger K, Giesinger JM, Kuster MS. All rights reserved. Version 1.1.

Appendix 1. Korean version of Forgotten Joint Score (FJS) Scoring: for scoring the FJS-12, all responses are summed (never, 0 points; almost never, 1 point; seldom, 2 points; sometimes, 3 points; mostly, 4 points) and then divided by the number of completed items. This mean value is subsequently multiplied by 25 to obtain a total score range of 0 to 100. Finally, the score is subtracted from 100 to change the direction of the final score in a way that a high score indicates a high degree of "forgetting" the artificial joint, that is, a low degree of awareness. If more than 4 responses are missing, the total score should not be used.