A New High-flexion Knee Scoring System to Eliminate the Ceiling Effect

Sang-Eun Na MD, Chul-Won Ha MD,
Choong-Hee Lee MD

Received: 20 January 2011 / Accepted: 15 November 2011 / Published online: 29 November 2011
© The Association of Bone and Joint Surgeons® 2011

Abstract

Background Various scoring systems document improvement after TKA, but most are associated with a ceiling effect that may fail to distinguish between patients having different levels of knee function after TKA. We therefore developed a new scoring system for patients with higher levels of flexion to eliminate ceiling effects observed with current systems.

Questions/purposes The purposes of this study were (1) to determine whether the high-flexion knee score eliminates the ceiling effect, (2) to assess the validity and responsiveness of the high-flexion knee score, and (3) to determine whether the high-flexion knee score can aid in differentiation of the knee status of patients at the ceiling level.

Methods We prospectively studied 165 patients with 201 well-functioning knees who had undergone primary TKA. We obtained Knee Society scores, WOMAC scores, Feller scores, SF-36 scores, and high-flexion knee scores for all patients. The high-flexion knee score includes items that

reflect knee function in the high functional range, such as sitting on or rising from the floor, squatting, or kneeling. We determined the ceiling effects and score distributions of various scoring systems. We performed a convergent validity test of the high-flexion knee score by correlation analysis with these various scoring systems. Responsiveness of the high-flexion knee score was assessed by correlation analysis of changes in various scoring systems. To determine whether the high-flexion knee score can aid in differentiation of knee status of patients at the ceiling level, relative responsiveness of the various scores in the ceiling versus below the ceiling range was determined.

Results The high-flexion knee score showed no ceiling effect, whereas the other systems did. Addition of the high-flexion knee score to the other scoring systems eliminated these ceiling effects and resulted in more normalized score distributions. The high-flexion knee score correlated ($r = -0.77$) with WOMAC in postoperative scores, and it also correlated with the changes in WOMAC ($r = -0.69$) and SF-36 physical functioning ($r = 0.62$). The correlation of WOMAC score with SF-36 physical function score was lower in patients at the ceiling level of the Knee Society knee score and Knee Society function score, compared with the correlation at below-ceiling range, whereas the high-flexion knee score maintained a correlation with the SF-36 physical function score, even at the ceiling level of the Knee Society knee score and Knee Society function score.

Conclusions Compared with other systems, the high-function knee score appears valid for evaluating the knee status in the high-flexion range. Our data suggest the high-flexion knee score differentiates among the knee status in the high-function range. Furthermore, the high-function knee score eliminates the ceiling effect of the currently used scoring tools, and thus may be useful when combined with other scoring systems.

Each author certifies that he has no 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.

Each author certifies that his institution approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research.

Electronic supplementary material The online version of this article (doi:10.1007/s11999-011-2203-5) contains supplementary material, which is available to authorized users.

S.-E. Na, C.-W. Ha (✉), C.-H. Lee
Department of Orthopedic Surgery, Samsung Medical Center,
Sungkyunkwan University School of Medicine, 50 Irwon-dong,
Gangnam-gu, Seoul 135-710, Republic of Korea
e-mail: hacw@skku.edu; chulwon.ha@gmail.com

Level of Evidence Level III, diagnostic study. See Guidelines for Authors for a complete description of levels of evidence.

Introduction

Various clinical scoring systems are used to assess patients after TKA [2, 7, 8, 12]. These scoring systems aggregate weighted scores for pain, ROM, stability, alignment, and functional ability. Using these systems, pain and function of the knee after TKA have been reported by numerous investigators [4, 6, 18, 24]. However, discrepancy exists between assessment by these scoring systems and patient satisfaction with the function or symptoms after TKA [14, 21]. Although patient satisfaction is a complex phenomenon affected by many variables [7, 14, 21], we believe the current scoring systems do not properly assess some aspects of knee status. For example, the Knee Society (KS) knee score [12] may not be sensitive enough to assess pain and function after the current TKAs, because, when using this system, most TKA scores are greater than 80 points at any time after surgery [12, 23, 29]. These similarly high scores suggest a ceiling effect, and these methods fail to differentiate between patients of more modest function from those who either expect or actually achieve higher levels of function after TKA, especially those who typically require more motion and/or greater strength and balance (ie, sitting on or rising from the floor, squatting for gardening or golfing, or kneeling for prayer, etc).

The KS knee score and function score assess only walking and stair climbing and fail to address more demanding physical activities, such as squatting, kneeling, or rising from the floor. For example, to achieve a maximum KS function score, patients need to be able to climb stairs and walk unlimited distances without the use of a walking aid, and active patients usually are able to meet these criteria. Accordingly, currently used KS scoring systems may be of little value for differentiation of functional status of patients in the high-function range who have had TKAs. The KS knee and function scores have not undergone formal validation and were developed on the basis of expert selection. The WOMAC score has been well validated, and addresses the activities of daily living, but ignores more demanding functions (ie, “high-function” ranges requiring greater motion or strength or balance). Furthermore, it also places less focus on more active patients in terms of delineating small functional differences in high-functional activities. Improvements in surgical technique and implant design have enabled some patients to regain knee function, even in the high-function range after TKA [9, 10, 15]. Further, with the improvements the indications have expanded to include younger patients with

expectations of high function and high flexion. Accordingly, we suspect the proportion of patients with high scores at approximately the ceiling level also has increased. We therefore developed a new scoring system for assessment of the knee in patients in the high-function range – the high-flexion knee score (HFKS) – to minimize the ceiling effect.

The purposes of this study were (1) to determine whether the HFKS eliminates the ceiling effect, (2) to assess the validity and responsiveness of the HFKS, and (3) to determine whether the HFKS can aid in differentiation of the knee status of patients at the ceiling level.

Patients and Methods

We prospectively followed 178 patients who underwent 214 posterior-stabilized fixed-bearing TKAs from August 2008 to February 2009. During that period, 204 patients underwent surgery, however, we excluded 11 with a history of knee infection and 15 who were unable to complete questionnaires owing to cognitive or language difficulties. Of the 178 eligible patients, 13 were lost to followup leaving 165 patients with 201 TKAs. The mean age of the patients was 68.7 years (range, 52–83 years). Diagnoses were osteoarthritis in 192 knees and rheumatoid arthritis in nine. This study protocol was approved by our institutional review board.

The HFKS questionnaire is composed of two pain items and seven function items (Table 1), which we selected based on our experience (The Korean version is available as a supplemental appendix with the online version of CORR). For pain, we addressed knee pain on ascending or descending stairs and knee pain on sitting on or rising from the floor. We asked patients to check on a five-point Likert scale for each item. Allocated scores were 5 (no pain), 4 (mild pain), 3 (moderate pain), 2 (severe pain), or 1 (extreme pain). The function questionnaire addressed performance regarding seven high functional activities: ascending stairs, descending stairs, rising from the floor, rising from a low chair, cross-legged sitting, squatting, and kneeling. In a similar manner, we asked patients to check on a five-point Likert scale for each item; the allocated scores were; 5 (no difficulty), 4 (mild difficulty), 3 (moderate difficulty), 2 (severe difficulty), or 1 (extreme difficulty). We then defined the HFKS as the sum of the pain and function scores. HFKS ranged from 9 to 45 points (45 points being the highest possible score). A high HFKS indicated a high level of performance, with the least pain or difficulty during high-functional activities.

All TKAs were performed by the same surgeon (CWH) using the same technique. In each case, the surgeon used anterior referencing intramedullary instruments in

Table 1. Items in the HFKS questionnaire

Activity	Score*				
High-flexion knee pain					
Ascending or descending stairs	No pain	Mild pain	Moderate pain	Severe pain	Extreme pain
Sitting on or rising from the floor	No pain	Mild pain	Moderate pain	Severe pain	Extreme pain
High-flexion knee function					
Ascending stairs	No difficulty	Mild difficulty	Moderate difficulty	Severe difficulty	Extreme difficulty
Descending stairs	No difficulty	Mild difficulty	Moderate difficulty	Severe difficulty	Extreme difficulty
Sitting on or rising from the floor	No difficulty	Mild difficulty	Moderate difficulty	Severe difficulty	Extreme difficulty
Sitting on or rising from a low chair	No difficulty	Mild difficulty	Moderate difficulty	Severe difficulty	Extreme difficulty
Cross-legged sitting	No difficulty	Mild difficulty	Moderate difficulty	Severe difficulty	Extreme difficulty
Squatting	No difficulty	Mild difficulty	Moderate difficulty	Severe difficulty	Extreme difficulty
Kneeling	No difficulty	Mild difficulty	Moderate difficulty	Severe difficulty	Extreme difficulty

HFKS = high-flexion knee score; *response to each item was scored as 5 (no pain or difficulty), 4 (mild), 3 (moderate), 2 (severe), or 1 (extreme).

preparation of the femur. The surgeon inserted an intra-medullary femoral guide at the distal femur and performed 5° valgus cutting with respect to the femoral anatomic axis. The femoral component then was inserted along the transepicondylar axis. In addition, the surgeon referenced the posterior condylar axis and AP axis of Whiteside and Arima [28]. The surgeon referenced alignment of the tibial component against the medial third of the tibial tubercle and checked patella tracking with trial components in place using the towel clip technique. The Genesis II prosthesis (Smith & Nephew, Memphis, TN, USA) was used in 62 knees, Scorpio-flex (Stryker, Mahwah, NJ, USA) in 43 knees, Triathlon (Stryker) in 39 knees, LPS-flex (Zimmer, Warsaw, IN, USA) in 37 knees, and Vanguard (Biomet, Warsaw, IN, USA) in 20 knees.

Patients underwent clinical assessments using the KS knee score and KS function score [12], WOMAC [2], the score proposed by Feller et al. [8], SF-36 [19], and the HFKS preoperatively, and postoperatively at 6 weeks, 3 months, 6 months, and 12 months, and annually thereafter. Data were collected by one independent, experienced research assistant (THS). The KS questionnaire includes items on functional ability, including walking distance, stair climbing ability, and use of walking aids. We used knee ROM, stability, alignment, and muscle power for calculation of KS clinical rating scores, which consist of separate knee and function scores ranging from 0 to 100 points (where 100 points is the best possible score). The KS function score allocates points for walking distance and stair-climbing ability and makes deductions for use of a walking aid; 100 points (the maximum allowed) represents an unlimited walking distance and normal stair-climbing ability without use of an aid. Regarding the KS knee score, we used Insall's modification made in 1993 [22], which allocates 8° for each point in ROM, thus we regarded the maximum point as 93, which represents a stable, well-aligned,

painless knee with 144° flexion. The WOMAC is a 24-item, self-administered health questionnaire specifically designed for patients with osteoarthritis of the knee or hip [2]. The WOMAC scoring system includes three categories that measure pain, joint stiffness, and physical function. WOMAC questions use a five-point Likert scale and scores are summed to obtain WOMAC scores. The maximum score possible is 96 points, which indicates a high level of difficulty or disability. Thus, we used an inverted scale of WOMAC to analyze the ceiling effect like other scoring systems. The questionnaire of Feller et al. [8] includes items on anterior knee pain, quadriceps strength, and ability to rise from a chair and climb stairs; these scores range from 3 to 30 points, with 30 points representing the best possible score. The SF-36 [18] includes a question asking patients to rate their current general health status compared with their status 12 months ago. Patients answer on a scale with five responses ranging from "much better" to "much worse." The standardized method of calculating the SF-36 domains was used, so that each of eight subgroups had a score of 0 to 100 points, with 100 points representing the best possible score. Therefore, the scales of these various scoring systems have different values (Table 2). The data set of this study cohort included complete preoperative and 12-month data. There were no missing data.

The ceiling is defined as the highest score of a scale, and the ceiling effect concerns the proportion of respondents who achieve the highest possible score [27]. To assess the presence of a ceiling effect, we calculated the percentage of patients who achieved the highest possible score for each of the five scoring systems, and investigated the score distributions using histograms and scatter plots.

Validity reflects the extent to which the instrument measures what it is purported to measure. Although the WOMAC and SF-36 are the only two well-validated tools

Table 2. Comparison of various scoring systems including in this study

Scoring system	Range of score		Number of items in the scoring system	Mean score from the cohort of this study
	Worst score	Best score		
SF-36 physical function [19]	0	100	10	62.9 (range, 20–90)
Knee Society knee score [22]	0	100	9	87.1 (range, 53.7–95)
Knee Society function score [22]	0	100	5	94.1 (range, 57–100)
Feller et al. [8]	3	30	4	27.6 (range, 18–30)
WOMAC [2]	96	0	24	17.4 (range, 4–38)
HFKS	9	45	9	30.4 (range, 24–39)

HFKS = High-flexion knee score.

Table 3. Ceiling effects of current scoring systems and changes in ceiling effects after adding HFKS

Scoring system	Ceiling effect*	Scoring system	Ceiling effect*
Knee Society knee score [22]	25%	Knee Society knee score + HFKS	0%
Knee Society function score [22]	43%	Knee Society function score + HFKS	0%
WOMAC [2]	0%	WOMAC + HFKS	0%
Feller et al. [8]	56%	Feller et al. + HFKS	0%
HFKS	0%		

HFKS = high-flexion knee score; * proportion of the highest score.

among the scoring systems included in this study, all of the included scoring systems have been used as reliable instruments in clinical research for measurement of knee status after TKA [3, 17]. Thus, we performed a convergent validity test of the HFKS by correlation analysis with these various scoring systems. Responsiveness of the HFKS was assessed by evaluation of whether changes in the scores of the HFKS showed correlation with changes in other scoring systems. To determine whether the HFKS can aid in differentiation of knee status of patients at the ceiling level, we dichotomized the cohort as two subgroups (patients at the ceiling level and those below the ceiling level), and compared the responsiveness of the WOMAC, SF-36 physical function, and HFKS among each other. We performed the Spearman rank correlation analysis using SAS 9.1.3 (SAS Institute, Cary NC, USA).

Results

The average scores of the five scoring systems were: KS knee score = 87.1 points (range, 53.7–95 points), KS function score = 94.1 points (range, 57–100 points), WOMAC = 17.4 points (range, 4–38 points), score of Feller et al. = 27.6 points (range, 18–30 points), HFKS = 30.4 points (range, 24–39 points), and SF-36 physical function = 62.9 points (range, 20–90 points) (Table 2).

The percentages of patients who achieved a maximum score (ie, ceiling) were: KS knee society = 25%, KS function score = 43%, WOMAC = 0%, score of Feller

et al. = 56%, and HFKS = 0%. However, the addition of HFKS to scores of the other scoring systems resulted in complete elimination of the ceiling effect for those systems (Table 3). The histogram of the distributions of postoperative scores obtained using the various scoring systems showed the HFKS produced a normal distribution pattern, whereas the distribution pattern of other scoring systems skewed toward the ceiling, particularly for the KS knee and function scores, and scores of Feller et al. (Fig. 1). However, after addition of the HFKS to the other scoring systems, the scoring systems showed a near normal distribution pattern as the ceiling effect was almost eliminated (Fig. 2). Scatter plots of the preoperative and postoperative scores of the HFKS and the various scoring systems clearly showed the ceiling effect of the KS knee and function scores, and score of Feller et al. as they level off over increasing HFKS, postoperatively (Fig. 3).

The convergent validity test revealed the strongest correlation ($r = -0.77$), of the HFKS with the WOMAC in postoperative scores, whereas the others showed weaker correlation, with the score of Feller et al. showing the weakest correlation ($r = 0.40$) (Table 4). Regarding the responsiveness of the HFKS, changes in the HFKS showed moderate correlation with the changes in the WOMAC and SF-36 physical function, whereas weak correlation was observed with the KS knee and function scores and score of Feller et al. (Table 5).

Subgroup analysis of patients at and below the ceiling levels revealed the correlation of WOMAC score and

Fig. 1A–E The histograms of the distributions of postoperative scores obtained using the (A) KS knee score and (B) KS function score show that the distribution pattern is skewed toward the ceiling. (C) The histogram of the distributions of postoperative scores obtained using the WOMAC index shows that the distribution pattern is more normalized than for the KSKS or KSFS, but still skewed toward the ceiling. (D) The histogram of the distributions of postoperative scores obtained using the score of Feller et al. [8] shows that the distribution pattern is skewed toward the ceiling. (E) The histogram of the distributions of postoperative scores obtained using the HFKS produced a normal distribution pattern. HFKS = high-flexion knee score; KSFS = Knee Society function score; KSKS = Knee Society knee score.

Fig. 2A–D The histograms after addition of the HFKS to the (A) KS knee score and (B) function score show a near-normal distribution pattern as the ceiling effect was almost eliminated. (C) The histograms after addition of the HFKS to the (C) WOMAC index and (D) score of Feller et al. show a near-normal distribution pattern as the ceiling effect was almost eliminated. HFKS = high-flexion knee score; KSFS = Knee Society function score; KSKS = Knee Society knee score.

Fig. 3A–E Scatter plots of the preoperative and postoperative scores of the HFKS and the (A) KS knee and (B) function scores show the ceiling effects of the KS knee and function scores as they level off over the increasing HFKS postoperatively. (C) Scatter plots of the preoperative and postoperative scores of the HFKS and the WOMAC index show no ceiling effect of the WOMAC index. (D) Scatter plots of the preoperative and postoperative scores of the HFKS and the score of Feller et al. show the ceiling effect of the score of Feller et al. as it levels off over the increasing HFKS postoperatively. (E) Scatter plots of the preoperative and postoperative scores of the HFKS and the SF-36 show no ceiling effect. HFKS = high-flexion knee score; KSFS = Knee Society function score; KSKS = Knee Society knee score.

SF-36 physical function score for patients at the ceiling level of the KS knee and function scores was reduced compared with the correlation at below the ceiling range, whereas the HFKS maintained moderate correlation with the SF-36 physical function score, even at the ceiling level of the KS knee and function scores (Table 6).

Discussion

Various scoring systems currently used show a ceiling effect as the functional status of knees improves after TKA. Thus, the current scoring systems might not differentiate between patients at the ceiling level with different levels of knee function after TKA. We therefore developed the HFKS to eliminate the ceiling effects with current scoring systems. Therefore, the purposes of this study were (1) to determine whether the HFKS eliminates the ceiling effect, (2) to assess the validity and responsiveness of the HFKS, and (3) to determine whether the HFKS can aid in differentiation of knee status of patients at the ceiling level.

The limitations of our study should be addressed. First, we included items in the HFKS based on our experience. Although the items were not selected through a formal item-generation or selection process, we believe the HFKS includes the most important items regarding the knee status in the high-flexion range. The fact that we found a correlation of the HFKS with the validated WOMAC score and SF-36 physical function score suggests the items selected were reasonable. Second, we assessed the validity of the HFKS by correlation analysis of the HFKS with other well-known or well-validated scoring systems. However, this method may not be sufficient for validating a new scoring system; therefore, additional studies for thorough validation of the HFKS are needed. Third, the HFKS did not incorporate the intensity and frequency of activities, and we did not evaluate the activity level of the patients with scales (eg, the activity scale developed by Tegner and Lysholm [26]). The HFKS is focused on evaluation of knee status; therefore, it can reflect only the status of the knee, and not the functional status of the patient. However, we found the HFKS eliminated the ceiling effect of well-known scoring systems, and had

Table 4. Results of correlation analyses among various scoring systems

	Preoperative						Postoperative					
	SF-36	KSKS	KSFS	Feller et al.	WOMAC	HFKS	SF-36	KSKS	KSFS	Feller et al.	WOMAC	HFKS
SF-36	1.00	0.35 (< 0.001)	0.24 (0.003)	0.18 (0.067)	-0.20 (0.022)	0.24 (0.004)	1.00	0.64 (< 0.001)	0.60 (< 0.001)	0.42 (< 0.001)	-0.38 (< 0.001)	0.55 (< 0.001)
KSKS		1.00	0.24 (< 0.001)	0.12 (0.083)	-0.52 (< 0.001)	0.47 (< 0.001)		1.00	0.42 (< 0.001)	0.26 (< 0.001)	-0.39 (< 0.001)	0.55 (< 0.001)
KSFS			1.00	0.06 (0.380)	-0.30 (< 0.001)	0.32 (< 0.001)			1.00	0.46 (< 0.001)	-0.77 (< 0.001)	0.74 (< 0.001)
Feller et al.				1.00	-0.04 (0.595)	0.09 (0.206)				1.00	-0.43 (< 0.001)	0.40 (< 0.001)
WOMAC					1.00	-0.66 (< 0.001)					1.00	-0.77 (< 0.001)
HFKS						1.00						1.00

KSKS = Knee Society knee score; KSFS = Knee Society function score; HFKS = high-flexion knee score.

Table 5. Responsiveness of various scoring systems

Scoring system	SF-36	Knee Society knee score	Knee Society function score	Feller et al.	WOMAC	HFKS
SF-36	1.00	0.62 (< 0.001)	0.55 (< 0.001)	0.40 (< 0.001)	-0.44 (< 0.001)	0.62 (< 0.001)
Knee Society knee score		1.00	0.20 (0.004)	0.17 (0.040)	-0.43 (< 0.001)	0.43 (< 0.001)
Knee Society function score			1.00	0.24 (0.002)	-0.52 (< 0.001)	0.43 (< 0.001)
Feller et al.				1.00	-0.27 (< 0.001)	0.30 (< 0.001)
WOMAC					1.00	-0.69 (< 0.001)
HFKS						1.00

HFKS = High-flexion knee score.

Table 6. Relative responsiveness of scoring systems in subgroups of patients at ceiling and below ceiling

Scoring system	Knee Society function score < 100			Knee Society knee score < 93.0			Knee Society knee score = 93.0			Feller et al. < 30			Feller et al. = 30		
	SF-36	WOMAC	HFKS	SF-36	WOMAC	HFKS	SF-36	WOMAC	HFKS	SF-36	WOMAC	HFKS	SF-36	WOMAC	HFKS
SF-36	1.00	-0.21 (0.024)	0.42 (< 0.001)	1.00	-0.44 (< 0.001)	0.54 (< 0.001)	1.00	-0.17 (0.221)	0.44 (0.001)	1.00	-0.37 (< 0.001)	0.45 (< 0.001)	1.00	-0.34 (< 0.001)	0.64 (< 0.001)
WOMAC		1.00	-0.59 (< 0.001)		1.00	-0.69 (< 0.001)		1.00	-0.57 (< 0.001)		1.00	-0.68 (< 0.001)		1.00	-0.60 (< 0.001)
HFKS			1.00			1.00			1.00			1.00			1.00

HFKS = high-flexion knee score.

good correlation and responsiveness with those systems. Therefore, we believe the HFKS is a novel tool for evaluation of knees in the high-function range. Fourth, we do not have data for the HFKS in a normal healthy population and the effect of aging of patients on the HFKS. Additional investigations of these topics are needed. Fifth, items in the HFKS include activities that are not commonly performed during activities of daily living in Western countries, such as kneeling or squatting. Accordingly, the percentage of patients at the ceiling level (maximum score) might be different or the ceiling level might be 80 to 90 points in Western populations. However, we believe the scoring system should include those activities for differentiation among knee functional status in the high-flexion range. Highly performing patients need their knees to allow high-flexion activities such as gardening, golfing, or praying, even in Western countries. We believe the difference in the capability of performing those activities is shown only by a scoring system like the HFKS. However, the relative importance of various features would likely differ from culture to culture and perhaps even within a given culture depending on various patient expectations.

We identified a ceiling effect of the current scoring systems, confirming previous reports suggesting most patients who have had TKAs score greater than 80 points at any time after surgery [12, 23, 29]. However, there is no reliable outcome measuring tool to specifically evaluate knee status before and after TKA that differentiates knee status in the ceiling range. We found the HFKS exhibited no ceiling effect, but rather yielded a wide range of scores around a relatively central mean. These observations suggest its usefulness as a scoring system. In addition, the distribution of the HFKS more closely follows a normal distribution than the other scoring systems, which award scores toward the high end of their scales. In our study the mean KS function score was 94.1 and scores fell in a narrow range at the top end of the scale. Regarding the difference between the KS function score and the HFKS, 86 knees (43%) scored 100 points. Of these 86 knees, 43 (50%) scored between 31 to 33 and 27 (31%) were between 34 to 36 in the HFKS, presumably because of some functional limitation or discomfort experienced during high-function activities like sitting cross-legged, squatting, or kneeling. Using current scoring tools, the difference in knee status in this high-function range cannot be shown. In these kinds of cases, the HFKS shows that knee function is suboptimal in the high-function range. Furthermore, although pain and function of the knee are much improved after TKA according to currently used outcome tools [1, 13, 16], patient satisfaction is only approximately 85%. We believe the use of the HFKS allows identification of some of the unsatisfactory aspects

of the outcome after TKA, which is a first step toward further enhancing the outcome of TKA. Furthermore, we found that adding the HFKS to the score of each of the other scoring systems eliminated the ceiling effect in each case and improved the score distributions; therefore, we believe that, in addition to being useful on its own, the HFKS could be used in conjunction with established scoring systems.

We confirmed the validity and responsiveness of the HFKS. As the convergent validity test revealed a correlation of the HFKS with the WOMAC in postoperative scores, and as the WOMAC and SF-36 were reported as more valid measures of outcomes of TKA than the KS knee and function scores [17], the HFKS appears to be a reasonable tool for evaluating the status of the knee after a TKA. The analysis of the responsiveness of the HFKS assessed by evaluation of whether changes in the scores of the HFKS showed correlation with changes in other scoring systems also revealed reasonable correlation between the change in HFKS and the changes in WOMAC and SF-36. Among a few different measures for evaluation of responsiveness, including the internal responsiveness statistics (paired t-test, effect size, or standardized responsiveness mean) and the external responsiveness statistics (area under the receiver operating characteristic [ROC] curve, correlation, or linear regression) [11, 20, 25, 30], the correlation method was used in the current study and also by Corzillius et al. [5] and Lingard et al. [17]. We believe the correlation method is appropriate for our study, because it reflects the extent to which changes in a measure over a specified time relate to corresponding changes in a reference measure of health status (SF-36 physical function in the current study), which is important to patients [11]. Therefore, this study suggests the HFKS will be a valid outcome measure for evaluation of knees after TKA.

Our data suggest the HFKS differentiates among the status of knees in the high-function range. The WOMAC score has been well validated and addresses activities of daily living, but it does not seem to consider knee status in the high-function range. Although the WOMAC score showed no ceiling effect, the subgroup analysis suggests better function of the HFKS in differentiating the knee status in the ceiling range.

The HFKS appears to be a valid scoring system for evaluating knee status in the high-flexion range. Our data suggest the HFKS differentiates knee status in the high-function range and eliminates the ceiling effect of the currently used scoring tools. It may be a useful tool when used in combination with other scoring systems.

Acknowledgments We thank Seonwoo Kim PhD and Minji Kim MS for support with statistical analysis and Tai-Hee Seo BS for efforts in collecting the raw data for this study.

References

1. Anderson JG, Wixson RL, Tsai D, Stulberg SD, Chang RW. Functional outcome and patient satisfaction in total knee patients over the age of 75. *J Arthroplasty*. 1996;11:831–840.
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. Bourne RB. Measuring tools for functional outcomes in total knee arthroplasty. *Clin Orthop Relat Res*. 2008;466:2634–2638.
4. Clayton RA, Amin AK, Gaston MS, Brenkel IJ. Five-year results of the Sigma total knee arthroplasty. *Knee*. 2006;13:359–364.
5. Corzillius M, Fortin P, Stucki G. Responsiveness and sensitivity to change of SLE disease activity measures. *Lupus*. 1999;8:655–659.
6. Cushnaghan J, Bennett J, Reading I, Croft P, Byng P, Cox K, Dieppe P, Coggon D, Cooper C. Long-term outcome following total knee arthroplasty: a controlled longitudinal study. *Ann Rheum Dis*. 2009;68:642–647.
7. Ethgen O, Bruyere O, Richy F, Dardennes C, Reginster JY. Health-related quality of life in total hip and total knee arthroplasty: a qualitative and systematic review of the literature. *J Bone Joint Surg Am*. 2004;86:963–974.
8. Feller JA, Bartlett RJ, Lang DM. Patellar resurfacing versus retention in total knee arthroplasty. *J Bone Joint Surg Br*. 1996;78:226–228.
9. Huang HT, Su JY, Wang GJ. The early results of high-flex total knee arthroplasty: a minimum of 2 years of follow-up. *J Arthroplasty*. 2005;20:674–679.
10. Huddleston JI, Scarborough DM, Goldvasser D, Freiberg AA, Malchau H. 2009 Marshall Urist Young Investigator Award: How often do patients with high-flex total knee arthroplasty use high flexion? *Clin Orthop Relat Res*. 2009;467:1898–1906.
11. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol*. 2000;53:459–468.
12. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. *Clin Orthop Relat Res*. 1989;248:13–14.
13. Jörn LP, Johnsson R, Toksvig-Larsen S. Patient satisfaction, function and return to work after knee arthroplasty. *Acta Orthop Scand*. 1999;70:343–347.
14. Kane RL, Maciejewski M, Finch M. The relationship of patient satisfaction with care and clinical outcomes. *Med Care*. 1997;35:714–730.
15. Kim TH, Lee DH, Bin SI. The NexGen LPS-flex to the knee prosthesis at a minimum of three years. *J Bone Joint Surg Br*. 2008;90:1304–1310.
16. Lieberman JR, Dorey F, Shekelle P, Schumacher L, Thomas BJ, Kilgus DJ, Finerman GA. Differences between patients' and physicians' evaluations of outcome after total hip arthroplasty. *J Bone Joint Surg Am*. 1996;78:835–838.
17. Lingard EA, Katz JN, Wright RJ, Wright EA, Sledge CB; Kinemax Outcomes Group. Validity and responsiveness of the Knee Society Clinical Rating System in comparison with the SF-36 and WOMAC. *J Bone Joint Surg Am*. 2001;83:1856–1864.
18. Loughhead JM, Malhan K, Mitchell SY, Pinder IM, McCaskie AW, Deehan DJ, Lingard EA. Outcome following knee arthroplasty beyond 15 years. *Knee*. 2008;15:85–90.
19. McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care*. 1993;31:247–263.

20. Moretti C, Viola S, Pistorio A, Magni-Manzoni S, Ruperto N, Martini A, Ravelli A. Relative responsiveness of condition specific and generic health status measures in juvenile idiopathic arthritis. *Ann Rheum Dis*. 2005;64:257–261.
21. Noble PC, Conditt MA, Cook KF, Mathis KB. The John Insall Award: Patient expectations affect satisfaction with total knee arthroplasty. *Clin Orthop Relat Res*. 2006;452:35–43.
22. Rationale of The Knee Society Clinical Rating System. Available at: <http://www.kneesociety.org/web/outcomes.html>. Accessed September 23, 2011.
23. Schai PA, Thornhill TS, Scott RD. Total knee arthroplasty with the PFC system. Results at a minimum of ten years and survivorship analysis. *J Bone Joint Surg Br*. 1998;80:850–858.
24. Shah AK, Celestin J, Parks ML, Levy RN. Long-term results of total joint arthroplasty in elderly patients who are frail. *Clin Orthop Relat Res*. 2004;425:106–109.
25. Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J Clin Epidemiol*. 1995;48:1369–1378.
26. Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. *Clin Orthop Relat Res*. 1985;198:43–49.
27. 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.
28. Whiteside LA, Arima J. The anteroposterior axis for femoral rotational alignment in valgus total knee arthroplasty. *Clin Orthop Relat Res*. 1995;321:168–172.
29. Wright J, Ewald FC, Walker PS, Thomas WH, Poss R, Sledge CB. Total knee arthroplasty with the kinematic prosthesis. Results after five to nine years: a follow-up note. *J Bone Joint Surg Am*. 1990;72:1003–1009.
30. Wright JG, Young NL. A comparison of different indices of responsiveness. *J Clin Epidemiol*. 1997;50:239–246.