

# Can Original Knee Society Scores Be Used to Estimate New 2011 Knee Society Scores?

Susan M. Odum PhD, Thomas K. Fehring MD,  
On Behalf of Knee Society Crosswalk Writing Group

Published online: 13 May 2016

© The Association of Bone and Joint Surgeons® 2016

## Abstract

**Background** The Knee Society Score (KSS) instrument is one of the most commonly reported primary outcome measures for total knee arthroplasty (TKA). Originally developed in 1989, the KSS was expanded and updated in 2011; however, the original KSS does not directly translate

into the 2011 KSS. To date, no conversion algorithm has been developed, hindering the ability of researchers to adopt the 2011 KSS while maintaining their historical/longitudinal original KSS data.

**Questions/purposes** The purpose of this study is to develop regression equations to map the original KSS to the 2011 KSS, allowing original and 2011 KSS data sets to be combined.

**Methods** In this multicenter, nonrandomized study, a convenience sample of 815 patients undergoing primary TKA completed the original KSS questionnaire and the 2011 KSS questionnaire. Additionally, patient gender, patient age, and patient ethnicity were recorded. These data were then used to generate regression models to estimate the 2011 objective and function KSS from the original KSS. Of the 815 study patients, 476 (58%) were female and 339 (42%) were male at an average age of 67 years (SD 9.4). Roughly half of patients were assessed preoperatively (430 of 815 [53%]) with the remaining patients assessed postoperatively (386 of 815 [47%]). The average followup for postoperative patients was 4.4 years (SD 3.5 years).

**Results** We have created a spreadsheet that can be used by individuals with no statistical training to crosswalk the objective and function subscores from the original KSS to the 2011 KSS [Supplemental materials are available with the online version of *CORR*®.]. The predictive model very accurately estimated the 2011 objective score, on average, within 0.22 points on the 100-point 2011 objective KSS at the cohort or aggregate level. The objective model accurately estimated the 2011 objective KSS within 8.83 points, on average, of the actual 2011 objective KSS at the individual patient level. However, as a result of large outliers, 37% of the estimated 2011 objective KSS were greater than 10 points from the actual 2011 objective KSS. To illustrate,

---

The institution of one or more of the authors (SMO, TKF) has received funding from the Knee Society, Clinical Research Grant. 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.

Each author certifies that his or her institution approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained.

Knee Society Crosswalk Writing Group members meeting authorship criteria: Doug Dennis MD, Colorado Joint Replacement; Raymond Kim MD, Colorado Joint Replacement; Stuart Goodman MD, PhD, Stanford University; James Huddleston MD, Stanford University; Greg Kazarian, Rothman Institute; Jess H. Lonner MD, Rothman Institute; Juan C. Suarez MD, Cleveland Clinic Florida; Preetesh Patel MD, Cleveland Clinic Florida; Giles R. Scuderi MD, Northshore LIJ Orthopedic Institute; Theo Zois PA, Northshore LIJ Orthopedic Institute; Scott Sporer MD, MS, Rush University and Northwestern Medicine Joint Replacement Institute Central DuPage Hospital; Thomas J. Blumenfeld MD, Joint Surgeons of Sacramento; Bryce A. Van Doren MPA, MPH, OrthoCarolina Research Institute; and Rebecca J. Haug BA, OrthoCarolina Research Institute.

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s11999-016-4886-0) contains supplementary material, which is available to authorized users.

---

S. M. Odum (✉)

OrthoCarolina Research Institute, Inc, 2001 Vail Avenue, Suite 300, Charlotte, NC 28207, USA  
e-mail: Susan.Odum@orthocarolina.com

T. K. Fehring

OrthoCarolina Hip & Knee Center, Charlotte, NC, USA

if you use the model to estimate the 2011 objective KSS on a cohort of 100 patients, a patient with an original objective KSS of 88 will have an estimated objective KSS between 79 and 97 points. On the other hand, if you calculate an average original objective KSS of 88 for all 100 patients, the estimated average 2011 objective KSS will be 88 for the group. The predictive model accurately estimated the 2011 function KSS within 0.14 points on the 1000-point 2011 function KSS at the cohort level. At the patient level, the 2011 function KSS was also estimated within 8.8 points of the actual 2011 function KSS. However, 43% of the estimated function scores were greater than 10 points of the actual 2011 function KSS.

**Conclusions** Clinicians and researchers can input their original KSS with demographic data into these equations to estimate the 2011 KSS objective and function scores. The small prediction error of 0.22 points that we calculated indicates that these models can be used to estimate the 2011 objective and function KSS at the aggregated cohort level. Although the average error score was within 10 points at the individual patient level, there was a high percentage of large errors resulting from outliers in the data set. These outliers seemed to be related to patients with excellent range of motion who had substantial pain and limited function or patients who have poor range of motion with excellent function and little pain. This may be inherent with the KSS or with the study sample. Nevertheless, one must use caution when estimating at the patient level. Additionally, the accuracy of the prediction scores decreases if any of the demographic variables included in this study are not available.

## Introduction

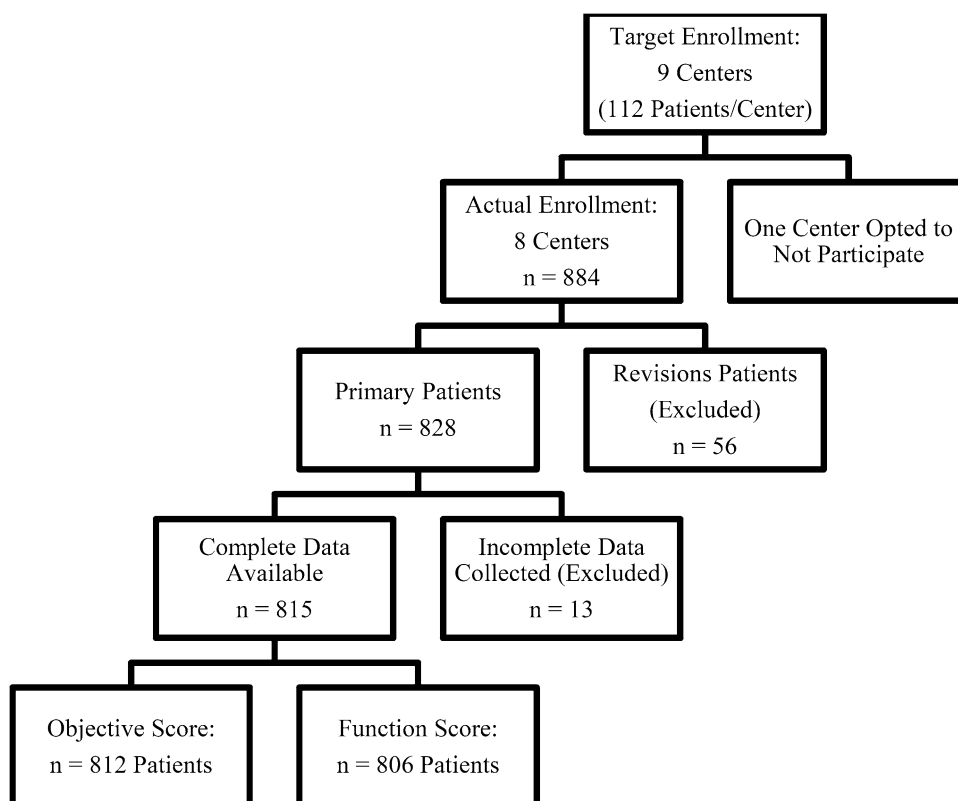
During the past few decades, several disease-specific measures have been used to evaluate outcomes after TKA and there has been wide variation in the reporting of these measures [14, 15, 18]. Such variation introduces challenges in comparing published studies, which can complicate clinical decision-making, research study design, and policy decisions. Riddle et al. [14] conducted a meta-analysis to determine the extent of the variation in the use of such measurement tools implemented in randomized clinical trials [14]. The findings indicate that the Knee Society Score (KSS) instrument is one of the most commonly reported outcome measures for knee arthroplasty [14]. Therefore, it is important to have the methodology and ability to use historical and longitudinal clinical data obtained using the original KSS while transitioning to the 2011 KSS.

The KSS was initially developed by the Knee Society in 1989 and further modified in 1993 [6]. Based on a panel of Knee Society surgical and clinical experts, the KSS was designed to have an objective knee score to evaluate the knee and a subjective patient function score. Having two separate scores avoids issues of degrading scores solely based on declining function resulting from the natural aging process. The objective knee score includes three main constructs: pain, knee stability, and ROM [6]. Of the maximum 100 points, a possible 50 points are assigned to pain, 25 points to stability, and 25 points for ROM with deductions for deformity and malalignment of the native joint and knee arthroplasty [6]. A well-aligned knee with no pain, 125° of motion, and good AP and mediolateral instability will achieve a KSS of 100 points [6]. The maximum function score is also 100 points and consists of patient-reported outcomes that measure walking distance (50 points) and the ability to ascend and descend stairs (50 points) with deductions for the use of assistive devices such as canes or walkers [6].

In 2009, the Knee Society Scoring Committee (KSSC) was formed as a multidisciplinary team of arthroplasty surgeons who were Knee Society members as well as epidemiologists and statisticians. Over a 3-year period, the KSSC developed and tested this 2011 KSS system using scientific, psychometric, and validation methodologies, which has been described by Noble et al. in great detail [9, 16]. The 2011 KSS expands on the original KSS and remains both provider- and patient-derived. The 2011 KSS includes an expanded functional component as well as new domains of patient satisfaction and patient expectations [10, 16]. Furthermore, there are minor changes in existing domains of pain, alignment, and stability. In addition to changes in the specific items, the point allocations are also different between the original and the 2011 KSS [6, 16].

A method to convert original KSS scores to 2011 KSS score has not been developed. As such, it is difficult to evaluate longitudinal clinical outcomes of TKA with the two similar, yet meaningfully different rating systems. The 2011 KSS instrument includes domains that are not covered in the original KSS, which presents a challenge to overcome without an acceptable crosswalk, or prediction, regression model. On the other hand, both versions of the KSS conceptually measure the constructs of pain, function, ROM, alignment, and stability. Given this conceptual similarity, we hypothesized that the regression models would crosswalk to relatively precise predictions, particularly with the addition of data that captured a broader clinical picture to improve the models' ability to estimate the 2011 KSS from the original KSS [9]. The specific aim of this study was to develop a regression equation that

**Fig. 1** Flow diagram illustrates how the study cohort was reached.



precisely converts the original KSS to the 2011 KSS to enable users to evaluate longitudinal outcomes after transitioning from the original KSS to the 2011 KSS.

## Patients and Methods

After institutional review board approval, preoperative and postoperative patients undergoing knee arthroplasty from eight high-volume orthopaedic clinics were recruited and enrolled in this study. In our multicenter, cross-sectional, nonrandomized study, a convenience sample of 828 patients undergoing primary TKA provided informed consent. Thirteen patients failed to complete all required forms and were excluded from subsequent analyses (Fig. 1). The final study cohort includes 815 study patients who completed the original KSS and 2011 KSS. Study patients were surveyed either pre- or postoperatively. Because it was critical to have a large variation in possible KSS and we were not interested in the change from pre- to postoperative, we did not collect pre- and post-data on the same patient. Additionally, patient gender, age, body mass index (BMI), and patient ethnicity were recorded. Data were entered and managed using the REDCap (Research Electronic Data Capture) data management system [5], which allowed for a multimodal data-capturing plan. In

other words, providers and patients could collect data on paper or enter it directly into the system.

## Study Sample Characteristics

Of the 815 study patients, 476 (58%) were female and 339 (42%) were male at an average age of 67 years (SD 9.4) (Table 1). The median BMI of the study group was 30 kg/m<sup>2</sup> (interquartile range, 27–35). The majority of the patients were white (681 of 815 [84%]). Roughly half of patients were assessed preoperatively (430 of 815 [53%]) with the remaining patients assessed postoperatively (386 of 815 [47%]). Nearly all of the preoperative patients were diagnosed with degenerative osteoarthritis (413 of 430 [96%]). The average followup for postoperative patients was 4.4 years (SD 3.5 years; range, 6 months to 8 years) with a total of 84% at greater than 2-year followup assumed to be fully recovered.

## Development of Prediction Models

As noted, the key objective was to develop regression models that allow for the prediction, or crosswalk, of the original KSS to the 2011 KSS. The independent variables for the

**Table 1.** Frequency and proportions of demographic characteristics for the primary TKA cohort

Demographic	Primary TKA (n = 815)	
	Frequency	Percent
Gender		
Female	476	58
Male	339	42
Visit		
Preoperative	430	53
Postoperative	385	47
Age group (years)		
<65	293	36
65–75	381	47
>75	141	17
Ethnicity		
White	681	84
Black	70	9
Hispanic	37	4
Other	27	3
Body mass index (kg/m <sup>2</sup> )		
<18.5	4	0.5
18.5–24.9	100	12
25.0–29.9	279	34
≥30	432	54
Followup (postoperative patients) (months)		
<6	27	7
6–24	31	8
24–60	203	53
60–120	91	24
≥120	26	7
Missing	5	1

regression models were selected using the components of the original KSS as well as patient demographic variables (age, race/ethnicity, BMI, and gender) and timing/surgical status (pre- versus postoperative) based on clinical knowledge. The distribution of the data (functional form) for each numeric independent variable was tested using a multivariable fractional polynomial (MFP) model. Additionally, the significance of each interaction term was tested for inclusion using MFP to determine whether there was a statistically significant impact of the independent variables on the estimated 2011 objective and function KSS. The MFP results indicated that each numeric independent variable followed a linear form and there were no interaction terms that had a significant effect on the 2011 KSS. Therefore, no squared or higher-order (nonlinear forms) terms or interaction terms were included in the regression models.

We also assessed the functional form, or distribution, of the dependent variables (2011 objective and function KSS) using the method proposed by Manning and Mullahy [6]. Based on the Manning and Mullahy method [6], the log-scale residuals were evaluated for the amount of kurtosis (skewness). The coefficient of kurtosis was  $> 3$ , which indicated that the distribution of both dependent variables (2011 objective and function KSS) had heavy-tailed log-scale residuals. Therefore, we log-transformed the 2011 KS objective and function KSS. Finally, one generalized estimating equation regression model was analyzed for each of the dependent variables.

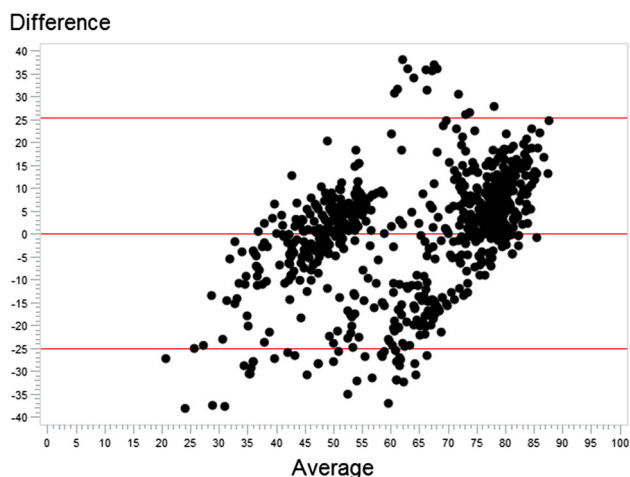
Predictive performance of the models was assessed using 10-fold crossvalidation. Essentially, 10 different, equally sized data sets were randomly constructed from our data and the error between the predicted 2011 KSS and the actual KSS was measured for each data set. These error measures were based on standard predictive performance criteria, known as forecasting errors, which measure the amount of error that exists between the actual (observed) 2011 KSS and the estimated 2011 KSS.

We included the following forecasting errors: mean absolute error (MAE), mean error (ME), and proportion of observations with errors of more than 10 points relative to the minimal clinically important difference (MCID). Although there is no MCID defined in the literature for the 2011 KSS, we estimated an MCID of 10 points ( $1/2$  of the calculated SD [3, 11–13]) using these data. These criteria are all negatively oriented in that lower values are better and reflect less error. The MAE is the average absolute error, meaning it quantifies the absolute difference between the actual score and the predicted score, or estimated score, at the individual patient level without regard for sign (ie, negative or positive number). The ME represents the average error when all of the error scores at the patient level are summed up and averaged across the study sample or cohort. Therefore, the ME is the difference between the actual score and the estimated score at the cohort level. Finally, we calculated the proportion of observations with absolute prediction errors for each model. We established a priori that the MAE would be the primary criterion used to recommend whether the models were appropriate to use to estimate 2011 KSS at the individual patient level. Moreover, we defined an MAE of 10 points (based on our estimated MCID of 10 points) as the cutoff for the acceptability of applying the model at the patient level. Predictive accuracy of the models was also evaluated using Bland-Altman plots, which provide a graphic illustration of the agreement between the observed KSS and the estimated KSS (Figs. 2, 3).

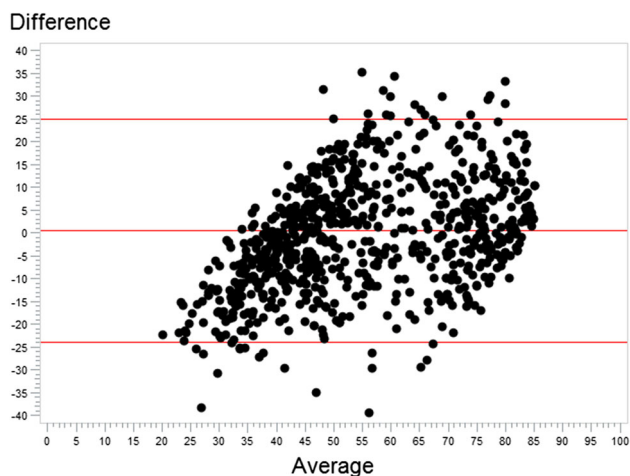
## Results

### KSS Conversion Tool and Illustrative Examples

We have provided a Microsoft Office Excel (Microsoft Inc, Redmond, WA, USA) worksheet that individuals can use with their original KSS to estimate the 2011 KSS (Appendix 1: Score Estimation Worksheet [Supplemental materials are available with the online version of *CORR*®.]). Users can enter their original KSS data into the worksheet, which will automatically estimate the 2011 objective and functional KSS. The worksheet also automatically calculates the mean and SD for the estimated



**Fig. 2** Bland-Altman plot depicts the average of the two methods (observed versus estimated 2011 objective KSS) by the difference of the two methods for our study cohort.



**Fig. 3** Bland-Altman plot depicts the average of the two methods (observed versus estimated 2011 function KSS) by the difference of the two methods for our study cohort.

scores at the cohort level. The spreadsheet can predict 2011 KSS for up to 100 patients at a time. Alternatively, researchers can use the coefficients from each regression model (Table 2) to manually calculate the 2011 objective and function KSS for individual patients. Importantly, when manually calculating individual scores, the regression output must be exponentiated ( $e^{\text{output}}$ ) to calculate the estimated score. The Score Estimation Worksheet automatically performs this step.

To explain how these predictive models estimate the 2011 KSS, we have provided illustrative examples. If one uses the model to estimate the 2011 objective KSS on a cohort of 100 patients, a patient with an original objective KSS of 88 will have an estimated objective KSS between 79 and 97 points. On the other hand, if you calculate an average original objective KSS of 88 for the entire cohort of all 100 patients, the estimated average 2011 objective KSS will be 88 for the group. Similarly, for an individual patient with an original function KSS of 60, the model will estimate a 2011 KSS between 51 and 69 on average. On the other hand, if the average original function KSS was 79, the model will estimate the 2011 function KSS to be 79 on average. More specific individual patient scenarios are illustrated subsequently.

#### Example Case 1

A 67-year-old black woman is seen preoperatively. Her BMI is 25.77 kg/m<sup>2</sup>. Her original objective KSS is 25, her original pain score is 20, her original ROM score is 18 (ROM = 90°), and her original function score is 50. When completing the 2011 KSS, her actual objective score is 68 points. Our crosswalk model predicted her 2011 objective KSS to be 69 points (1-point difference). When completing the 2011 KSS, her actual function score is 41 points. Our crosswalk model predicted her 2011 function score to be 39 points (2-point difference).

#### Example Case 2

A 59-year-old Hispanic man is seen postoperatively. His BMI is 44.51 kg/m<sup>2</sup>. His original objective KSS is 25 points, his original pain score is 10 points, his original ROM score is 23 points (ROM = 115°), and his original function score is 35 points. When completing the 2011 KSS, his actual objective score is 73 points. Our crosswalk model predicted his 2011 objective KSS to be 80 points (7-point difference). When completing the 2011 KSS, his actual function score is 33 points. Our crosswalk model predicted his 2011 function score to be 37 points (4-point difference).

**Table 2.** Coefficients required for converting the original Knee Society Scores to new 2011 Knee Society Scores

2011 objective score		2011 function score	
Variable	Coefficient	Variable	Coefficient
Intercept (Constant)	3.4333720	Intercept (Constant)	3.2112340
Original objective score	0.0192571	Original function score	0.0086701
Original pain score	-0.0014589	Original pain score	0.0059676
Original ROM score	0.0118821	Original ROM score	0.0065693
Age	0.0001238	Age	-0.0011247
Body mass index	0.0038847	Body mass index	-0.0046646
Female	0	Female	0
Male	-0.0082558	Male	-0.0128300
Preoperative	0	Preoperative	0
Postoperative	0.0642019	Postoperative	0.2248505
White	0	White	0
Black	0.0268367	Black	-0.0107303
Hispanic	-0.0268751	Hispanic	-0.0399797
Other race	-0.0545634	Other race	-0.0549654

To estimate the 2011 KSS from the original KSS, you must exponentiate the output of the above equation (ie,  $e^{\text{[sum of equation]}}$ ); KSS = Knee Society Score.

### Example Case 3

A 64-year-old white woman is seen preoperatively. Her BMI is 27.47 kg/m<sup>2</sup>. Her original objective KSS is 0, her original pain score is 30, her original ROM score is 24 (ROM = 115°), and her original function score is 60. When completing the 2011 KSS, her actual objective score is 29 points. Our crosswalk model predicted her 2011 objective KSS to be 44 points (15-point difference). When completing the 2011 KSS, her actual function score was 58 points. Our crosswalk model predicted her 2011 function score to be 48 points (10-point difference).

### Detailed Analysis

The regression model developed to estimate, or predict, the 2011 objective KSS fit the data fairly well with a pseudo R<sup>2</sup> of 0.60, which means that the model explained 60% of the variability of the estimated 2011 objective KSS. This model that estimated the 2011 objective KSS was more accurate at the aggregate, or cohort level, than it was estimating at the individual patient level. On average, at the cohort level, the model estimated the 2011 objective score within 0.22 points (ME = 0.22) on the 100-point 2011 objective KSS. At the individual patient level, the model estimated the 2011 objective KSS within 8.83 points (MAE = 8.83), on average, of the actual 2011 objective KSS. Although this MAE of 8.83 is fairly accurate, there were some wide misses in the data. A total of 37% of the

estimated 2011 objective KSS was greater than 10 points from the actual 2011 objective KSS.

The accuracy of the regression model developed to estimate the 2011 function KSS was similar in accuracy as the model that predicted the 2011 objective KSS. The model explained a total of 68% (pseudo R<sup>2</sup> = 0.68) of the variability in the estimated 2011 function KSS. The predictive model accurately estimated the 2011 function KSS within 0.14 points (ME = 0.14) on the 100-point 2011 function KSS at the cohort level. At the patient level, the 2011 function KSS was also estimated within 8.80 points (MAE = 8.80) of the actual 2011 function KSS. However, 43% of the estimated function scores were greater than 10 points of the actual 2011 function KSS.

The wide misses, with greater than 35% of the observations that have estimated scores that are not within 10 points of the actual scores, seem to occur with patients who had incongruent ROM with other Knee Society subscores. In other words, patients who had excellent ROM but poor scores on other aspects of the KSS such as alignment, pain, and function were prone to an inaccurate estimate of the 2011 objective KSS. Similarly, the estimated 2011 objective KSS was inaccurate for patients with poor ROM and good scores on other subscores.

### Discussion

As a commonly used outcome measure in TKA studies, the new 2011 KSS was a welcome revision with added

functionality compared with the original KSS developed in 1989. However, researchers struggled with how to incorporate the new 2011 KSS when considering years of data collected earlier using the original KSS. The purpose of this study was to develop a regression model that can be used to crosswalk the original KSS to the 2011 KSS to enable users to evaluate longitudinal outcomes after transitioning from the original KSS to the 2011 KSS. We developed one model to estimate the 2011 objective KSS and one model to estimate the 2011 function KSS.

The limitation of the study is inherent to the methodology. In a review of crosswalk studies and methodology, Mortimer and Segal [8] addressed the methodological and conceptual concerns with regression crosswalk methodology and found large variation in the explanatory power of regression models. A broad overview of study results indicates that the explanatory power of the regression models is lower when converting scores with disparate constructs and measurement scales [6]. However, both versions of the KSS conceptually measure the constructs of pain, function, ROM, alignment, and stability. The similarity of pain and function domains allowed us to derive precise estimates of the objective and function subscores with good explanatory power of both models (pseudo  $R^2$  of 0.60 for objective and 0.68 for function). However, the additional 2011 KSS domains of patient expectations and patient satisfaction were not included. Therefore, we can only offer crosswalk algorithms for the objective and function subscores. These data also represent private clinics and academic medical centers with a varied patient population. It is thus reasonable to expect that these data and models can be generalized for use with any TKA patient population. Although these models are not perfect, the relatively small prediction errors that we report indicate that the models can be used to estimate the 2011 objective and function KSS at the aggregate, cohort level. At the individual level, the MAEs are less than the MCID, indicating that these models can be applied at the individual level; however, caution should be exercised when using the models in outliers or patients with conflicting subscores (eg, high ROM but poor performance on other subscores). Certainly, if any of the data variables are not accessible, we caution use of either model at the individual patient level as the prediction errors increase without the robust data.

Much of this type of mapping, or crosswalk, methodology has been previously conducted using other outcome measurement tools with the purpose of predicting a generic, preference-based healthy utility measure from a disease-specific measure or two generic health-related quality-of-life instruments [2, 4, 8]. To our knowledge, there are no equivalent studies in arthroplasty that offer a direct comparison to our results. Although the use of regression models to map, or crosswalk, values from disease-specific health

measurement instruments to utilities derived from generic health measurement instruments is increasing, there have been few studies related to osteoarthritis and total joint arthroplasty [1, 4, 8, 9, 17]. Furthermore, the WOMAC is the only disease-specific instrument that has been mapped to predict a preference-based health utility. Grootendorst et al. [3] developed a model that robustly converts the WOMAC to the Health Utilities Index 3 (HUI3) within 20 points in patients undergoing TKA and Marshall et al. [7] validated that model in patients undergoing THA with equally good results. Similarly, Barton et al. [1] estimated the EuroQol 5 Dimension (EQ5D) from the WOMAC within 18 points. In comparison, our predictive criteria values indicate smaller errors, that is, lower error values, in estimating 2011 KSS from the original (1989) KSS.

In conclusion, the results of this study offer a fairly precise algorithm that can crosswalk the original KSS objective scores and KSS function scores to the 2011 KSS objective and function scores. The low predictive errors indicate that these algorithms can be used at the individual patient level. However, if any of the variables in the regression model that we developed and are proposing are missing in the original score, our recommendation is to only use the scores at the aggregate level (ie, average values for a large cohort). We have made available the regression models and coefficients that the user can select and use based on their available data. Furthermore, a Microsoft Office Excel program is provided (Appendix 1: Score Estimation Worksheet [Supplemental materials are available with the online version of *CORR*®.]) that clinicians and researchers can input their historical KSS with demographic data into to crosswalk to the 2011 KSS objective and function scores.

**Acknowledgments** We thank Judy Guerreiro, Stefanie Miller, Matt Austin MD, Jaime Wanner, Bozena Kuczynski PA-C, Angela Bye, Kathy Hwang, William J. Maloney MD, Rose Walter, Todd Miner MD, Andrea Hankins, Ahmed Attia MD, and Leticia Barksdale MD, for assistance with study coordination and patient recruitment. In addition, the authors wish to thank Christi Cadd, Caryn Thompson, and Amy Roznowski of the OrthoCarolina Research Institute for their logistical support throughout the study. The authors also wish to thank Walter Beaver MD, William Griffin MD, J. Bohannon Mason MD, Brian Curtin MD, Bryan Springer MD of the OrthoCarolina Hip & Knee Center for recruiting patients for the study. Finally, the authors express their gratitude to the Knee Society for awarding the 2013 Knee Society Research Grant to support this study.

## References

1. Barton GR, Sach TH, Avery AJ, Doherty M, Jenkinson C, Muir KR. Comparing the performance of the EQ-5D and SF-6D when measuring the benefits of alleviating knee pain. *Cost Eff Resour Alloc.* 2009;7:12.
2. Farivar SS, Liu H, Hays RD. Half standard deviation estimate of the minimally important difference in HRQOL scores? *Expert Rev Pharmacoecon Outcomes Res.* 2004;4:515–523.

3. Grootendorst P, Marshall D, Pericak D, Bellamy N, Feeny D, Torrance GW. A model to estimate health utilities index mark 3 utility scores from WOMAC index scores in patients with osteoarthritis of the knee. *J Rheumatol*. 2007;34:534–542.
4. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42:377–381.
5. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. *Clin Orthop Relat Res*. 1989;248:13–14.
6. Manning WG, Mullahy J. Estimating log models: to transform or not to transform? *J Health Econ*. 2001;20:461–494.
7. Marshall D, Pericak D, Grootendorst P, Gooch K, Faris P, Frank C, Bellamy N, Torrance G, Feeny D. Validation of a prediction model to estimate health utilities index Mark 3 utility scores from WOMAC index scores in patients with osteoarthritis of the hip. *Value Health*. 2008;11:470–477.
8. Mortimer D, Segal L. Comparing the incomparable? A systematic review of competing techniques for converting descriptive measures of health status into QALY-weights. *Med Decis Making*. 2008;28:66–89.
9. Noble PC, Scuderi GR, Brekke AC, Sikorskii A, Benjamin JB, Lonner JH, Chadha P, Daylamani DA, Scott WN, Bourne RB. Development of a new Knee Society scoring system. *Clin Orthop Relat Res*. 2012;470:20–32.
10. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41:582–592.
11. Norman GR, Sloan JA, Wyrwich KW. The truly remarkable universality of half a standard deviation: confirmation through another look. *Expert Rev Pharmacoecon Outcomes Res*. 2004;4:581–585.
12. Rai SK, Yazdany J, Fortin PR, Avina-Zubieta JA. Approaches for estimating minimal clinically important differences in systemic lupus erythematosus. *Arthritis Res Ther*. 2015;17:143.
13. Riddle DL, Stratford PW, Bowman DH. Findings of extensive variation in the types of outcome measures used in hip and knee replacement clinical trials: a systematic review. *Arthritis Rheum*. 2008;59:876–883.
14. Riddle DL, Stratford PW, Singh JA, Strand CV. Variation in outcome measures in hip and knee arthroplasty clinical trials: a proposed approach to achieving consensus. *J Rheumatol*. 2009;36:2050–2056.
15. Scuderi GR, Bourne RB, Noble PC, Benjamin JB, Lonner JH, Scott WN. The new Knee Society Knee Scoring System. *Clin Orthop Relat Res*. 2012;470:3–19.
16. Segal L, Day SE, Chapman AB, Osborne RH. Can we reduce disease burden from osteoarthritis? *Med J Aust*. 2004;180:S11–17.
17. Singh J, Sloan JA, Johanson NA. Challenges with health-related quality of life assessment in arthroplasty patients: problems and solutions. *J Am Acad Orthop Surg*. 2010;18:72–82.
18. Singh J, Sloan JA, Johanson NA. Challenges with health-related quality of life assessment in arthroplasty patients: problems and solutions. *J Am Acad Orthop Surg*. 2010;18:72–82.