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

CrossMark

SYMPOSIUM: 2016 KNEE SOCIETY PROCEEDINGS

Can Preoperative Patient-reported Outcome Measures Be Used to Predict Meaningful Improvement in Function After TKA?

**Jonathan L. Berliner MD, Dane J. Brodke BA, Vanessa Chan MPH,
Nelson F. SooHoo MD, Kevin J. Bozic MD, MBA**

Abstract

Background Despite the overall effectiveness of total knee arthroplasty (TKA), a subset of patients do not experience expected improvements in pain, physical function, and quality of life as documented by patient-reported outcome measures (PROMs), which assess a patient's physical and emotional health and pain. It is therefore important to develop preoperative tools capable of identifying patients unlikely to improve by a clinically important margin after surgery.

J. L. Berliner, D. J. Brodke, V. Chan
Department of Orthopaedic Surgery,
University of California San Francisco,
San Francisco, CA, USA

V. Chan
Philip R. Lee Institute for Health Policy Studies,
University of California San Francisco,
San Francisco, CA, USA

Questions/purposes The purpose of this study was to determine if an association exists between preoperative PROM scores and patients' likelihood of experiencing a clinically meaningful change in function 1 year after TKA.
Methods A retrospective study design was used to evaluate preoperative and 1-year postoperative Knee injury and Osteoarthritis Outcome Score (KOOS) and SF-12 version 2 (SF12v2) scores from 562 patients who underwent primary unilateral TKA. This cohort represented 75% of the 750 patients who underwent surgery during that time period; a total of 188 others (25%) either did not complete PROM scores at the designated times or were lost to follow-up. Minimum clinically important differences (MCIDs) were calculated for each PROM using a distribution-based method and were used to define meaningful clinical improvement. MCID values for KOOS and SF12v2 physical component summary (PCS) scores were calculated to be 10 and 5, respectively. A receiver operating characteristic analysis was used to determine threshold values for preoperative KOOS and SF12v2 PCS scores and their respective predictive abilities. Threshold values defined the point after which the likelihood of clinically meaningful improvement began to diminish. Multivariate regression was used to control for the effect of preoperative mental and emotional health, patient attributes quantified by SF12v2 mental component summary (MCS) scores, on

N. F. SooHoo
Department of Orthopaedic Surgery, University of California
Los Angeles, Santa Monica, CA, USA

K. J. Bozic (✉)
Department of Surgery and Perioperative Care, Dell Medical
School at the University of Texas at Austin, 1400 Barbara Jordan
Boulevard, Suite 1.114, Austin, TX 78723, USA
e-mail: kevin.bozic@austin.utexas.edu

 Springer

patients' likelihood of experiencing meaningful improvement in function after surgery.

Results   Threshold values for preoperative KOOS and SF12v2 PCS scores were a maximum of 58 (area under the curve [AUC], 0.76; p < 0.001) and 34 (AUC, 0.65; p < 0.001), respectively. Patients scoring above these thresholds, indicating better preoperative function, were less likely to experience a clinically meaningful improvement in function after TKA. When accounting for mental and emotional health with a multivariate analysis, the predictive ability of both KOOS and SF12v2 PCS threshold values improved (AUCs increased to 0.80 and 0.71, respectively). Better preoperative mental and emotional health, as reflected by a higher MCS score, resulted in higher threshold values for KOOS and SF12v2 PCS.

Conclusions   We identified preoperative PROM threshold values that are associated with clinically meaningful improvements in functional outcome after TKA. Patients with preoperative KOOS or SF12v2 PCS scores above the defined threshold values have a diminishing probability of experiencing clinically meaningful improvement after TKA. Patients with worse baseline mental and emotional health (as defined by SF12v2 MCS score) have a lower probability of experiencing clinically important levels of functional improvement after surgery. The results of this study are directly applicable to patient-centered informed decision-making tools and may be used to facilitate discussions with patients regarding the expected benefit after TKA.

Level of Evidence   Level III, prognostic study.

Introduction

Despite the overall effectiveness of TKA, a subset of patients experience unsatisfactory results with respect to pain, function, and restoration of quality of life [39, 45]. For a patient to ultimately benefit from a surgical intervention, the procedure must result in a meaningful improvement in health, pain, or function such that the patient might consider repeating the intervention. The minimum clinically important difference (MCID) can be used to specify this degree of change in health and is one way to define what constitutes a successful outcome after a surgical intervention [8, 18, 23]. Prior studies that have defined and evaluated the MCID after TKA have shown that a notable proportion of patients do not experience this degree of change postoperatively [9, 11]. Furthermore, numerous studies have documented dissatisfaction rates after TKA ranging from 15% to 20% [25, 39]. A recent appropriateness study of patients from the Osteoarthritis Initiative concluded that 34% of TKAs were

performed inappropriately [36]. Regional, racial, and gender variations in patient selection throughout the United States also highlight the need to better define surgical appropriateness criteria and develop improved shared decision-making tools [14].

Improved quality of life and functional ability are often considered the most important outcomes after major joint replacement, and patients' own assessment of these outcomes is a key element in evaluating the effectiveness of the procedure. Consequently, the focus of outcomes assessment has shifted to a more patient-centered approach with the use of patient-reported outcome measures (PROMs). Disease-specific measures such as the Knee injury and Osteoarthritis Outcome Score (KOOS) are more sensitive to change within the context of a specific illness. Generic measures such as the SF-36 and SF-12 version 2 (SF12v2) provide information on overall health status, incorporate elements of psychosocial health, and have the ability to capture the effects of medical comorbidities on quality of life [21].

There is much interest in the ability to preoperatively identify patients who are at greatest risk of unsatisfactory outcomes after TKA as well as those who will benefit most. Preoperative pain and functional status, as measured by PROMs, have been shown to predict pain and functional ability after TKA [15, 27, 40]. More specifically, although patients with higher levels of preoperative pain and disability demonstrate the greatest improvements in PROM scores, they do not achieve absolute postoperative scores comparable to patients with less preoperative pain and better baseline function [13, 24]. A substantial body of evidence has also demonstrated that mental and emotional health influence postoperative PROMs after TKA [4, 5, 13, 16, 17, 19, 24, 26, 27, 40, 43, 44]. Worse preoperative mental and emotional health is associated with smaller improvements in physical function scores after TKA [4, 5, 16, 17, 22, 27, 40]. These findings highlight the importance of considering both the physical and mental components of preoperative PROMs if they are to be incorporated into a decision-making tool. To our knowledge, this has not been done for patients undergoing TKA.

The purpose of this study was to determine the association between preoperative PROM scores and the likelihood that patients undergoing TKA would experience meaningful clinical improvement 1 year after surgery, as defined by the MCID. Specifically, we asked whether (1) calculated threshold values would be acceptably predictive and define maximum preoperative functional component scores after which the likelihood of experiencing a meaningful improvement begins to diminish; and (2) controlling for baseline mental and emotional health would improve the predictive ability of these threshold values.

Patients and Methods

Data included in this study were obtained from a longitudinally maintained joint replacement outcomes registry, including multiple surgeons, from the primary author's institution (JLB). The database includes clinical information and patient-reported outcomes for patients undergoing TKA. Patient-reported outcomes were collected preoperatively and 1 year after surgery as part of routine care. The database also includes patient demographic information including age, gender, and race. Patients selected for this study had a history of primary unilateral TKA and available PROM data recorded at both preoperative and 1-year postoperative time points. To ensure the analysis was performed on a homogenous patient population, the data analysis excluded patients with pathological fracture, malignant neoplasm, or a history of a subsequent procedure on the knee. It was anticipated that 500 patients would be needed to establish appropriateness thresholds, allowing the assessment of receiver operating characteristic (ROC) curves and areas under the curve (AUCs) with a standard error of no more than 0.03 for an expected AUC of 0.7.

In total, 562 patients with knee osteoarthritis who underwent primary TKA between 2009 and 2013 met our inclusion criteria. Average patient age was 67 years (SD ± 11), 59% of the patients included in the study were women, and 76% were white. This cohort of patients represented 75% of the 750 patients undergoing primary, unilateral TKA included in our institution's joint replacement registry who had no history of a subsequent procedure. The remaining 188 patients were excluded from the study because they either did not have 1-year followup PROM scores available (60 patients) or were lost to follow-up entirely (128 patients). No differences were found between the study cohort and all patients lost to follow-up when comparing preoperative and postoperative PROM scores (Table 1).

Preoperative and 1-year postoperative KOOS and SF-12v2 PROMs were collected through an electronic interface or on paper by a research assistant (DP). The KOOS consists of 42 items separated into five subscales and is scored from 0 to 100 with 0 being the worst level of pain and function. The SF12v2 is a revised version of the original SF-12 with wording modifications to improve readability and ease of use. SF12v2 physical and mental composite scores (PCS and MCS, respectively) range from 0 to 100 in which a score of 0 indicates the lowest level of health. The scores of both subscales are calculated from the survey's 12 questions. Survey questions that assess mental and emotional health address the effect of emotional problems (such as feeling depressed or anxious) on work or regular daily activities. PROM scores and SDs were calculated using the scoring algorithms for each outcomes instrument. The SF12v2 PCS and MCS components were considered as separate outcome measures and not as individual subscales.

MCID may be calculated using consensus, anchor, or distribution-based methods [23, 28]. The MCID after TKA has been defined using the WOMAC and SF-36 PROMs and may be estimated, by using a distribution-based method, as half the SD of outcome change scores for a given instrument [11, 30, 46]. This method of calculation was chosen after evaluating several other possible estimates of the MCID, including distribution-based methods that use either 95% confidence intervals or half the interquartile range (IQR)/2. MCIDs were calculated separately for KOOS and SF12v2 PCS as half the SD of all change scores for that specific PROM [30]. Anchor-based methods require a separate subjective assessment measure of a patient's perceived benefit from an intervention, data not collected by our institution's joint replacement registry [47]. The calculated MCID value was 10 for the KOOS and 5 for the SF12v2 PCS (Table 2). Overall, 82% of patients achieved improvement greater than the MCID on the KOOS and 69% on the SF12v2 PCS after unilateral primary TKA.

A nonparametric ROC analysis was used to determine an optimal threshold value for KOOS and SF12v2 PCS separately. For each PROM, the calculated threshold value specified a preoperative score best able to predict the likelihood of a patient experiencing a MCID after TKA. The Youden index was applied to each PROM's ROC analysis to calculate threshold values [48]. The Youden index was used to define threshold values by calculating the preoperative PROM score with the highest combined sensitivity and specificity that a patient would experience the MCID. The c-statistic (AUC) of each ROC analysis indicated the predictive validity of this binary classifier test for predicting the likelihood that a patient would achieve the MCID. The c-statistic ranges from 0 to 1 with 1 indicating perfect discrimination and 0.5 indicating no better than chance. Predictive models are considered reasonable when the c-statistic (AUC) is > 0.7 and excellent when > 0.8 [20]. For this study, c-statistics > 0.7 were considered acceptably predictive.

Given the known effect of a patient's preoperative mental and emotional health on the functional result of TKA, a multivariate analysis was performed. A two-stage hierarchical multivariate logistic regression analysis was performed to determine the relative influence of preoperative MCS score on patients' likelihood of achieving a MCID based on their preoperative KOOS or PCS score. This analysis was necessary to control and adjust for individual patient preoperative variability to allow for comparisons between patients' clinically meaningful improvements. New Youden thresholds for KOOS and

Table 1. Comparison of study cohort to patients without 1-year PROM data

Group characteristic	Study cohort	Missing 1-year data	p value[*]
Number of patients	562	188	
Age (years)[†]	67 (11)	64 (11)	0.005
White	76%	67%	0.058
Female	60%	57%	0.617
Preoperative KOOS	50 (18)	48 (19)	0.229
Preoperative SF12v2 PCS	32 (10)	32 (10)	0.573
Preoperative SF12v2 MCS	52 (12)	51 (11)	0.700
Postoperative KOOS	80 (18)	80 (20)[‡]	0.948
Postoperative SF12v2 PCS	42 (12)	40 (11)[‡]	0.137
Postoperative SF12v2 MCS	52 (10)	53 (10)[‡]	0.639

* To compare the mean PROM scores of the study cohort with the mean PROM scores of those without 1-year data, a Student's t-test was used and a Fisher's exact test was used for categorical data; [†]mean (SD); [‡]includes patients with at least one postoperative PROM score available (but missing 1-year PROM data). This group represented 32% (n = 60) of the total patients not included in the study cohort. Postoperative PROM scores included in the mean were those available from latest followup. This ranged from 2 to 4 years postoperatively; PROM = patient-reported outcome measures; KOOS = Knee injury and Osteoarthritis Outcome Score; SF12v2 = SF-12 version 2; PCS = physical component summary; MCS = mental component summary.

Table 2. Threshold values for univariate and multivariate analysis

PROM	Thresholds	AUC	Sensitivity	Specificity	p value
KOOS	< 58	0.76 (0.03)	73% (4%)	69% (4%)	< 0.001
SF12v2 PCS	< 34	0.65 (0.04)	66% (4%)	56% (5%)	< 0.001
Multivariate SF12v2 MCS/KOOS[*]	See Fig. 2A	0.80	82%	63%	0.004
Multivariate SF12v2 MCS/SF12v2 PCS[*]	See Fig. 2B	0.71	64%	69%	< 0.001

* Also controlled for gender, age, race; 95% confidence intervals reported in parentheses; PROM = patient-reported outcome measure; AUC = area under the curve; KOOS = Knee injury and Osteoarthritis Outcome Score; SF12v2 = SF-12 version 2; PCS = physical component summary; MCS = mental component summary.

PCS were then calculated from the fitted logistic regression equation of the predicted probability of obtaining the MCID, generating a new threshold value for each potential preoperative MCS score. These new threshold values were then used to calculate new c-statistics to determine changes in the predictive ability of the KOOS and PCS threshold values after controlling for preoperative MCS scores.

Results

The calculated threshold values for functional outcome measures KOOS and SF12v2 PCS were 58 and 34, respectively. These threshold values defined the maximum preoperative scores after which the likelihood of a patient experiencing a minimum clinically important difference began to diminish. The KOOS threshold value of 58 proved to be acceptably predictive of a patient's likelihood of achieving the MCID with an AUC value of 0.76 (Fig. 1A). The SF12v2 PCS threshold value did not demonstrate sufficient predictive ability with an AUC of 0.65 (Fig. 1B).

The corresponding sensitivity and specificity values for each threshold ranged from 56% to 82% (Table 2).

Multivariate analysis, adjusting for preoperative mental and emotional health, demonstrated that patients with better baseline mental and emotional health (higher preoperative MCS scores) and the greatest amount of disability (lower KOOS and PCS scores) had the highest likelihood of experiencing a clinically meaningful improvement in function after TKA. In addition, the predictive ability of both the KOOS and SF12v2 PCS threshold values improved after controlling for mental and emotional health. The KOOS c-statistic improved from 0.76 to 0.80 and the SF12v2 PCS c-statistic improved from 0.65 to 0.71 (Table 2). Only when taking into consideration a patient's preoperative mental health are the SF12v2 PCS threshold values acceptably predictive. Using the fitted logistic regression equation, new threshold values for KOOS and SF12v2 PCS were calculated for each potential preoperative MCS value from 0 to 100. Higher preoperative SF12v2 MCS scores resulted in higher threshold values for both KOOS (Fig. 2A) and SF12v2 PCS (Fig. 2
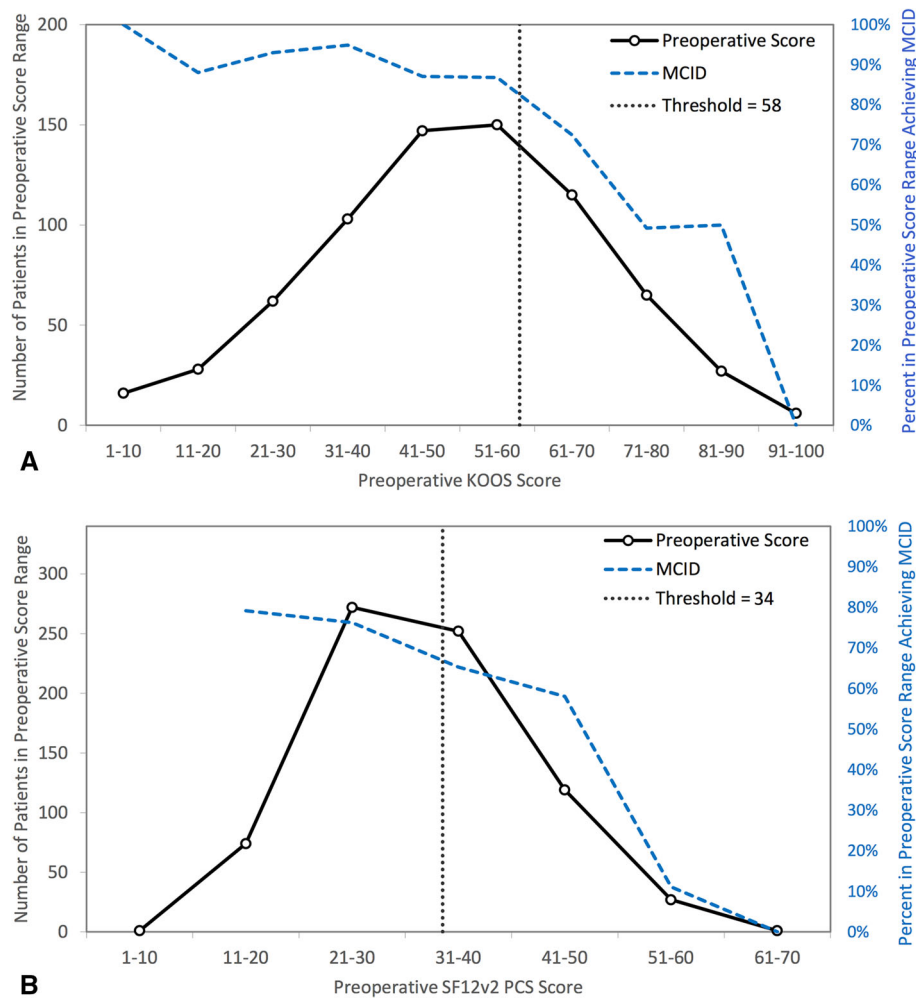
Fig. 1A–B The calculated threshold values, indicated by the dotted vertical lines, do not represent true cutoffs but instead serve to represent points after which a patient's likelihood of experiencing a clinically meaningful improvement in function begins to more rapidly diminish. (**A**) The KOOS threshold value of 58 was acceptably predictive of a patient's likelihood of experiencing a clinically meaningful improvement in outcome as measured by the 1-year postoperative HOOS score (AUC, 0.76). (**B**) The SF12v2 PCS threshold value of 34 was not acceptably predictive of a patient's likelihood of experiencing a clinically meaningful improvement in outcome as measured by the 1-year postoperative SF12v2 PCS score (AUC, 0.65).

B) such that for each 10-point increase in preoperative SF12v2 MCS score, there was an approximate 6-point increase in both KOOS and SF12v2 PCS threshold values. These findings suggest that patients with better mental and emotional health are more likely to experience a clinically meaningful improvement in function after TKA despite having superior baseline function.

Discussion

Despite the proven effectiveness of TKA, a notable subset of patients does not experience meaningful clinical improvement after surgery. Prior studies aimed at defining the MCID after TKA have demonstrated that 12% to 51%

of patients do not experience this degree of improvement postoperatively with respect to pain and function [2, 7, 9, 11]. With limited appropriateness criteria, the decision to pursue surgery is complex and multifactorial for both the patient and physician. This is evidenced by reports that suggest up to 34% of TKAs are performed inappropriately [3, 33, 34, 36]. Prior studies have attempted to define explicit clinical criteria for TKA appropriateness; however, the subjective nature of the procedure's indications requires patients to weigh the risks and benefits on the basis of their own values [10, 12]. This study is the first of which we are aware to identify an association between baseline functional status, adjusted for mental and emotional health, and the subset of patients most likely to experience a clinically meaningful improvement in function after TKA.
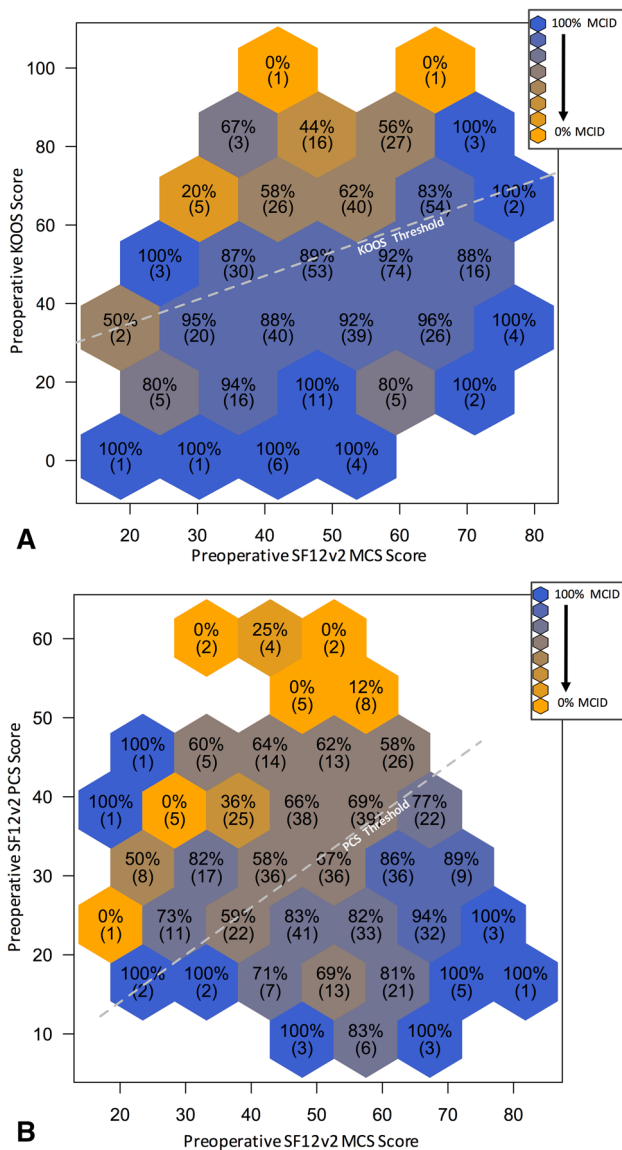
Fig. 2A–B SF12v2 PCS and KOOS threshold values (represented by dashed lines) are dependent on preoperative MCS score and demonstrate a linear relationship. Postoperative data are plotted in a binned fashion, which demonstrates the likelihood of attaining a MCID across different preoperative PROM score combinations. Hexagonal cells are labeled and shaded according to the proportion of patients within that cell who obtained the MCID (absolute number of patients in parentheses). By situating patients within a specific bin, one is able to visualize an approximate likelihood of obtaining a MCID based on preoperative PROM scores in the context of calculated threshold values. (A) The predictive ability of KOOS threshold values improved from 0.76 to 0.80 after adjusting for preoperative mental and emotional health. (B) After adjusting for preoperative mental and emotional health, SF12v2 PCS threshold values demonstrated an improved predictive ability (AUC, 0.71).

The preoperative threshold value for KOOS, determined to be 58, is capable, with acceptable predictive ability, of differentiating patients more likely to experience meaningful improvements in function after TKA from those who

are not. Additionally, threshold values for KOOS and SF12v2 PCS vary and demonstrate improved predictive ability when taking into account preoperative mental and emotional health.

This study has several limitations. The definition of a "successful" outcome is a controversial issue and varies between different patients and providers. For the purpose of this study it was defined as a change in PROM score after TKA greater than the MCID, which may not be the most appropriate definition of success. This definition excludes patient satisfaction, a separate outcome that has been shown to be dependent on patient expectations and highly variable as a result of factors specific to each individual patient [31, 37]. For this reason, patient satisfaction may be an unreliable metric to assess the impact of an intervention [37]. Additionally, the method of threshold value calculation using Youden's index, which maximizes the combined sensitivity and specificity of the cutoff point, may not be the most clinically relevant method. This may explain the relatively high proportion of patients in our study that fall outside of the defined thresholds when compared with prior studies of TKA use and appropriateness based on clinical criteria [10, 34]. Surgeons may prefer thresholds with higher sensitivity at the expense of specificity, thereby predicting that a greater percentage of patients are likely to experience meaningful functional benefit after TKA. Furthermore, this retrospective study does not confirm the ability of threshold values to predict which patients will actually experience a clinically meaningful improvement in practice. Only future prospective studies are able to validate the clinical application of threshold values. For these reasons, the calculated threshold values should not be regarded as true appropriate use criteria, but rather as tools to enhance patient education and shared decision-making.

Values for MCID can vary substantially and are dependent on numerous factors including the method of calculation, differing patient populations, and length of patient followup. For instance, prior studies designed to estimate the MCID for WOMAC and SF-36 after TKA using anchor-based methods have found consistently higher values than those using distribution-based methods [7, 11, 42]. This may suggest that distribution-based methods underestimate the amount of postoperative improvement necessary to be meaningful for patients. However, an ideal means of calculating MCID with regard to TKA, or any intervention, remains to be determined [7, 9, 11]. The minimum clinically important difference for a specific intervention is ultimately defined by what is interpreted as important to a patient and is therefore not a fixed attribute. This study used a distribution-based method that, although widely used, is generally not a preferred method as a result of several limitations. Distribution-based

methods are founded entirely on statistical reasoning and therefore fail to incorporate patients' own assessment of their condition. In the current study, we performed several different distribution-based estimates of MCID to evaluate for statistical variability. MCID was calculated using three different methodologies including half SD, 95% confidence intervals, and half the IQR. The half the SD method was chosen because it fell between the other proposed MCID values (Supplemental Table 1 [Supplemental materials are available with the online version of CORR®.]). Additionally, the 95% confidence intervals were very wide, whereas the IQR/2 criterion resulted in MCID values attained by a relatively small percentage of the patient population. Furthermore, every attempt was made to control for individual variability using multivariate techniques.

Similar to other regional and national joint replacement registries in the United States, our institution's joint replacement registry does not collect subjective patient assessments of improvement after TKA (other than as reported in their patient-reported outcome scores). We were therefore unable to perform an anchor-based method of calculation. However, because MCID values are known to be sample-specific, we favored a method that used data from our study population over adopting MCID values defined in previous studies. Applicable to any study that uses MCID, the reader should be made aware of the associated limitations and the resulting impact on clinical applications.

Length of followup for this study (1 year) may be considered a limitation. However, we believe that 1 year of follow-up is appropriate given the objective of our study and evidence from prior studies related to time to full recovery after TKA. When quantified with PROMs, the greatest change with regard to pain, function, and mental health has been shown to occur within the first 6 months after TKA and plateaus 1 year postoperatively [15, 35, 38]. One hundred eighty-eight (25%) of the patients from our institution's joint replacement registry who met the original inclusion criteria during the years 2009 to 2013 were not included in the study. This included 128 who were lost to follow-up and 60 who did have postoperative follow-up but not at the 1-year time point. Importantly, no differences were found between the study cohort and all patients lost to follow-up when comparing preoperative and postoperative PROM scores (Table 1). Given the distribution-based method used in this study, the value of MCID is dependent on variability within our patient population and therefore could be affected by the 128 patients lost to follow-up entirely. If 1-year postoperative PROM scores for those lost to follow-up were biased as compared with our study population, either more or less improved, our calculated MCID values would be larger. Although this would result in a smaller percentage of patients within our study population experiencing clinically meaningful improvement

after TKA, it would be unlikely to significantly affect threshold values, and corresponding c-statistics, because these are objective measures.

This study was performed at a single institution on a predominantly white, North American population. Accordingly, the results may not be applicable to individuals who are underrepresented in our patient population. However, both the KOOS and SF12v2 have been shown to have good applicability across varying populations and we therefore believe that our findings can be generalized. In the future, we believe that a methodology similar to the one described in this study may be applied to surgeon-specific data with the use of a computational algorithm. Such an algorithm could be incorporated into joint replacement registries as an application capable of generating surgeon-specific threshold values. Given the progressive adoption of regional and national joint replacement registries, this type of application would have limited barriers to entry and broad-reaching implications.

The threshold value for KOOS, which was a maximum of 58 out of a possible 100 points, was sufficiently predictive of attaining a MCID (AUC, 0.76). This suggests that patients with baseline KOOS scores greater than 58 are progressively less likely to experience a clinically meaningful improvement after surgery. This trend of diminishing returns with higher baseline functional capacity has been previously described [24, 26]. These findings are consistent with prior evidence suggesting that preoperative pain and functional status are predictive of functional ability after TKA [15, 27, 40]. Comparatively, the threshold value for SF12v2 PCS was not acceptably predictive. This is likely explained by the fact that generic PROMs such as the SF12v2 are less sensitive to changes in health after TKA when compared with disease-specific PROMs such as the KOOS [29].

The predictive ability of both SF12v2 PCS and KOOS threshold values improved after controlling for baseline mental and emotional health, as quantified by preoperative SF12v2 MCS scores. In fact, only when taking into consideration patients' preoperative mental and emotional health do the SF12v2 PCS threshold values become acceptably predictive. Additionally, the multivariate analysis demonstrated that baseline SF12v2 MCS scores paralleled functional threshold values. These findings are consistent with prior evidence, which demonstrates that poorer baseline mental and emotional health is associated with smaller improvement in function after TKA [4, 5, 13, 16, 17, 19, 24, 26, 27, 40, 43, 44]. By comparison, patient comorbidities and age have little effect on PROM scores after TKA [13, 32].

For physicians, the results of this study broaden the application of widely used patient-reported outcome surveys by providing preoperative threshold values for a

disease-specific PROM, KOOS, and a generic PROM, SF12v2, that have been adjusted for mental and emotional health. These data may help to identify the subset of patients with preoperative PROM scores that place them at a low likelihood of experiencing a clinically meaningful benefit after TKA. This type of information may facilitate further discussions surrounding the timing of surgery or the need for additional preoperative interventions before proceeding to surgery. More specifically, patients with low MCS scores may benefit from preoperative interventions aimed at improving mental and emotional health such as a multimodal program including cognitive therapy and education to better align patient expectations with realistic outcomes. Patients with high preoperative function may be counseled that the incremental improvement they can expect after surgery may not be sufficient to make a meaningful impact on their quality of life and therefore may not be considered worth the risk of undergoing a major surgical procedure. For patients, this information could help to give them a sense of what to expect after surgery, thereby promoting further engagement in the decision-making process. Evidence-based tools such as decision aids have been shown to be an effective means of helping patients make difficult decisions and lead to better outcomes and patient satisfaction after joint replacement surgery [1, 6, 41, 49]. Future prospective studies are needed to assess the use of preoperative PROM threshold values in shared clinical decision-making between patients with advanced knee osteoarthritis and their care team as well as their ability to increase the percentage of patients who experience a clinically meaningful improvement after TKA.

References

1. Adam JA, Khaw F-M, Thomson RG, Gregg PJ, Llewellyn-Thomas HA. Patient decision aids in joint replacement surgery: a literature review and an opinion survey of consultant orthopaedic surgeons. *Ann R Coll Surg Engl.* 2008;90:198–207.

2. Alzahrani K, Gandhi R, Debeer J, Petruccelli D, Mahomed N. Prevalence of clinically significant improvement following total knee replacement. *J Rheumatol.* 2011;38:753–759.

3. Ang DC, James G, Stump TE. Clinical appropriateness and not race predicted referral for joint arthroplasty. *Arthritis Care Res.* 2009;61:1677–1685.

4. Ayers DC, Franklin PD, Ploutz-Snyder R, Boisvert CB. Total knee replacement outcome and coexisting physical and emotional illness. *Clin Orthop Relat Res.* 2005;440:157–161.

5. Ayers DC, Franklin PD, Trief PM, Ploutz-Snyder R, Freund D. Psychological attributes of preoperative total joint replacement patients. *J Arthroplasty.* 2004;19:125–130.

6. Bozic KJ, Chiu V. Emerging ideas: Shared decision making in patients with osteoarthritis of the hip and knee. *Clin Orthop Relat Res.* 2011;469:2081–2085.

7. Chesworth BM, Mahomed NN, Bourne RB, Davis AM. Willingness to go through surgery again validated the WOMAC clinically important difference from THR/TKR surgery. *J Clin Epidemiol.* 2008;61:907–918.

8. Copay AG, Subach BR, Glassman SD, Polly DW, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J.* 2007;7:541–546.

9. Escobar A, García Pérez L, Herrera-Espiñeira C, Aizpuru F, Sarasqueta C, Gonzalez Sáenz de Tejada M, Quintana JM, Bilbao A. Total knee replacement; minimal clinically important differences and responders. *Osteoarthritis Cartilage.* 2013;21:2006–2012.

10. Escobar A, Quintana JM, Aróstegui I, Azkárate J, Güenaga JI, Arenza JC, Garai I. Development of explicit criteria for total knee replacement. *Int J Technol Assess Health Care.* 2003;19:57–70.

11. Escobar A, Quintana JM, Bilbao A, Aróstegui I, Lafuente I, Vidaurreta I. Responsiveness and clinically important differences for the WOMAC and SF-36 after total knee replacement. *Osteoarthritis Cartilage.* 2007;15:273–280.

12. Escobar A, Quintana JM, Bilbao A, Ibañez B, Arenaza JC, Gutiérrez L, Azkárate J, Güenaga JI, Vidaurreta I. Development of explicit criteria for prioritization of hip and knee replacement. *J Eval Clin Pract.* 2007;13:429–434.

13. 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.

14. Fisher ES, Bell J-E, Tomek IM, Esty AR, Goodman DC. *Trends and Regional Variation in Hip, Knee, and Shoulder Replacement. Dartmouth Atlas Surgery Report.* Hanover, NH, USA: The Dartmouth Institute for Health Policy and Clinical Practice; 2010.

15. Fortin PR, Clarke E, Joseph L, Liang MH, Tanzer M, Ferland D, Phillips C, Partridge J, Bélisle P, Fossel H, Mahomed N, Sledge CB, Katz JN. Outcomes of total hip and knee replacement: preoperative functional status predicts outcomes at six months after surgery. *Arthritis Rheum.* 1999;42:1722–1728.

16. Franklin PD, Li W, Ayers DC. The Chitranjan Ranawat Award: functional outcome after total knee replacement varies with patient attributes. *Clin Orthop Relat Res.* 2008;466:2597–2604.

17. Gandhi R, Dhotar H, Razak F, Tso P, Davey JR, Mahomed NN. Predicting the longer term outcomes of total knee arthroplasty. *Knee.* 2010;17:15–18.

18. Gatchel RJ, Mayer TG. Testing minimal clinically important difference: consensus or conundrum? *Spine J.* 2010;10:321–327.

19. Giesinger J, Kuster M. Association of psychological status and patient-reported physical outcome measures in joint arthroplasty: a lack of divergent validity. *Qual Life Outcomes.* 2013;11:64.

20. Gortmaker SL, Hosmer DW, Lemeshow S. Applied logistic regression. *Contemp Sociol.* 1994;23:159.

21. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med.* 1993;118:622–629.

22. Heck DA, Robinson RL, Partridge CM, Lubitz RM, Freund DA. Patient outcomes after knee replacement. *Clin Orthop Relat Res.* 1998;356:93–110.

23. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials.* 1989;10:407–415.

24. Jones C, Voaklander D, Suarez-Almazor M. Determinants of function after total knee arthroplasty. *Phys Ther.* 2003:696–706.

25. Judge A, Arden NK, Cooper C, Kassim Javaid M, Carr AJ, Field RE, Dieppe P. Predictors of outcomes of total knee replacement surgery. *Rheumatology (Oxford).* 2012;51:1804–1813.

26. Kahn TL, Soheili A, Schwarzkopf R. Outcomes of total knee arthroplasty in relation to preoperative patient-reported and radiographic measures: data from the Osteoarthritis Initiative. *Geriatr Orthop Surg Rehabil.* 2013;4:117–126.

27. Lingard E, Katz JN, Wright E, Sledge CB. Predicting the outcome of total knee arthroplasty. *J Bone Joint Surg Am.* 2004;86: 2179–2186.

28. McGlothlin A, Lewis R. Minimal clinically important difference: defining what really matters to patients. *JAMA.* 2014;312:1342–1343.

29. McGuigan FX, Hozack WJ, Moriarty L, Eng K, Rothman RH. Predicting quality-of-life outcomes following total joint arthroplasty. Limitations of the SF-36 Health Status Questionnaire. *J Arthroplasty.* 1995;10:742–747.

30. 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.

31. Palazzo C, Jourdan C, Descamps S, Nizard R, Hamadouche M, Anract P, Boisgard S, Galvin M, Ravaud P, Poiraudeau S. Determinants of satisfaction 1 year after total hip arthroplasty: the role of expectations fulfilment. *BMC Musculoskelet Disord.* 2014;15:53.

32. Peter WF, Dekker J, Tilbury C, Tordoir RL, Verdegaal SHM, Onstenk R, Bénard MR, Vehmeijer SB, Fiocco M, Vermeulen HM, van der Linden-van der Zwaag HMJ, Nelissen RGHH, Vliet Vlieland TPM. The association between comorbidities and pain, physical function and quality of life following hip and knee arthroplasty. *Rheumatol Int.* 2015:1233–1241.

33. Quintana JM, Arostegui I, Escobar A, Azkarate J, Goenaga JI, Lafuente I. Prevalence of knee and hip osteoarthritis and the appropriateness of joint replacement in an older population. *Arch Intern Med.* 2008;168:1576–1584.

34. Quintana JM, Escobar A, Arostegui I, Bilbao A, Azkarate J, Goenaga JI, Arenaza JC. Health-related quality of life and appropriateness of knee or hip joint replacement. *Arch Intern Med.* 2006;166:220–226.

35. Rat AC, Guillemin F, Osnowycz G, Delagoutte JP, Cuny C, Mainard D, Baumann C. Total hip or knee replacement for osteoarthritis: mid- and long-term quality of life. *Arthritis Care Res.* 2010;62:54–62.

36. Riddle D, Jiranek W, Hayes C. Using a validated algorithm to judge the approriateness of total knee arthroplasty in the United States: a multi-center longitudinal cohort study. *Arthritis Rheumatol.* 2014;66:2134-2143.

37. Ring D, Leopold SS. Editorial—Measuring satisfaction: can it be done? *Clin Orthop Relat Res.* 2015;473:3071–3073.

38. Rissanen P, Aro S, Sintonen H, Slätis P, Paavolainen P. Quality of life and functional ability in hip and knee replacements: a prospective study. *Qual Life Res.* 1996;5:56–64.

39. Scott CEH, Howie CR, MacDonald D, Biant LC. Predicting dissatisfaction following total knee replacement: a prospective study of 1217 patients. *J Bone Joint Surg Br.* 2010;92:1253–1258.

40. Sharma L, Sinacore J, Daugherty C, Kuesis DT, Stulberg SD, Lewis M, Baumann G, Chang RW. Prognostic factors for functional outcome of total knee replacement: a prospective study. *J Gerontol A Biol Sci Med Sci.* 1996;51:M152–157.

41. Slover J, Alvarado C, Nelson C. Shared decision making in total joint replacement. *JBJS Rev.* 2014;2:e1–e6.

42. SooHoo NF, Li Z, Chenok KE, Bozic KJ. Responsiveness of patient reported outcome measures in total joint arthroplasty patients. *J Arthroplasty.* 2015;30:176–191.

43. Sullivan M, Tanzer M, Stanish W, Fallaha M, Keefe FJ, Simmonds M, Dunbar M. Psychological determinants of problematic outcomes following total knee arthroplasty. *Pain.* 2009;143:123–129.

44. Utrillas-Compaired A, De La Torre-Escuredo BJ, Tebar-Martínez AJ, Asúnsolo-Del Barco Á. Does preoperative psychologic distress influence pain, function, and quality of life after TKA? *Clin Orthop Relat Res.* 2014;472:2457–2465.

45. Wylde V, Dieppe P, Hewlett S, Learmonth ID. Total knee replacement: is it really an effective procedure for all? *Knee.* 2007;14:417–423.

46. Wyrwich KW, Bullinger M, Aaronson N, Hays RD, Patrick DL, Symonds T. Estimating clinically significant differences in quality of life outcomes. *Qual Life Res.* 2005;14:285–295.

47. Wyrwich KW, Norquist JM, Lenderking WR, Acaster S. Methods for interpreting change over time in patient-reported outcome measures. *Qual Life Res.* 2013;22:475–483.

48. Youden WJ. Index for rating diagnostic tests. *Cancer.* 1950;3:32–35.

49. Youm J, Chenok K, Belkora J, Chan V, Bozic KJ. The emerging case for shared decision making in orthopaedics. *J Bone Joint Surg Am.* 2012;94:1907–1912.