Published in final edited form as:

J Arthroplasty. 2016 November ; 31(11): 2402–2407.e2. doi:10.1016/j.arth.2016.04.015.

Pre-operative Pain and Function: Profiles of Patients Selected for Total Knee Replacement

Uyen-Sa D. T. Nguyen, DSc, MPH, David C. Ayers, MD, Wenjun Li, PhD, Leslie R. Harrold, MD, MPH, and Patricia D. Franklin, MD, MBA, MPH

University of Massachusetts Medical School, Department of Orthopedics and Physical Rehabilitation, 55 Lake Ave North, Worcester, MA 01655

Abstract

Background—Total knee replacement (TKR) is an effective treatment to relieve pain and restore function in patients with advanced knee osteoarthritis (OA). TKR utilization is growing rapidly, and the appropriateness of current TKR use is of great interest. We examined patient-reported pre-operative pain and function profiles to understand symptom severity at the time of TKR decision.

Methods—Data were from FORCE-TJR. We included patients undergoing primary, unilateral TKRs between 2011 and 2014 for OA, and had data on the Knee Injury and Osteoarthritis Outcome Score (KOOS) pain and SF-36 Physical Component Summary (PCS) score. We compared patient profiles across groupings by symptoms: 1) little pain-and-high function (KOOS ≥ 70, PCS ≥ 40), 2) little pain-but-poor function (KOOS ≥ 70, PCS < 40), 3) high pain-but-high function (KOOS < 70, PCS ≥ 40), and 4) high pain-and-poor function (KOOS < 70, PCS < 40).

Results—Of 6,936 patients, 77% had high pain-and-poor function (Group 4), 19% had high pain or poor function (Groups 2–3), and 5% had little pain-and-high function prior to TKR (Group 1). In Group 1, 86% were constantly aware of their knee problem, 48% reported pain daily yet 5% experienced severe or extreme pain-on-stairs, 1% pain-in-bed. Over half had a lot of limitations in vigorous activities. Compared with Group 4, Group 1 were older, less obese, more educated, and included more men and people reporting being healthy, less disabled, and happy ($p < 0.05$ for all).

Conclusion—The majority of patients undergoing TKRs had significant pain and/or poor function. Our results provide critical information given the current debate of potentially inappropriate TKR utilization in the US.

Keywords

Total Knee Replacement; Total Knee Arthroplasty; Pain; Function

Corresponding Author: Uyen-Sa D. T. Nguyen, DSc, MPH Address: 55 Lake Ave North, Worcester, MA 01655, Phone: 508-856-1611; uyensa.nguyen@umassmed.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of Interest Statement

The authors have no financial or personal relationships with other people or organizations that could bias our current research.

INTRODUCTION

When knee arthritis pain is severe and frequent, and limits mobility and daily activities, a total knee replacement (TKR) remains the most effective treatment to relieve arthritis pain and to improve function¹. In the US, the annual rate of TKRs among people 65 years or older increased almost 9-fold between 1979 and 2006²⁻⁴. In addition, the latest US hospital-discharge data indicate a significant increase in TKR utilization among the younger patients (< 65 years of age)⁵. Among the 719,000 TKRs performed in 2010, close to 50% were performed in those under 65 years of age. While likely multifactorial, the reasons for the increased TKR utilization remain unclear.

Both the increasing demand and associated costs have resulted in a surge in interest regarding the appropriateness of TKR as surgical interventions. Escobar⁶ determined appropriateness of TKR timing and use based on the patient's age, radiographic evidence of severity and localization of arthritic disease, knee joint mobility and stability, knee pain and function status, and a prior history of surgical and non-surgical treatments. Using a modified version of Escobar's criteria, a study recently found that close to 30% of TKR patients in the Osteoarthritis Initiative (OAI) did not meet the "appropriate" surgical intervention criteria⁷. Whether these appropriateness criteria are culturally suitable and appropriate for current patients requiring joint replacement remain unclear. It is possible that a more relevant appropriateness criteria would balance patient-reported data with clinical data⁸.

The purpose of our current study was to ascertain the characteristic of patients selected for TKR in a recently established US nationally-representative cohort of TJR patients. Our goals were: 1) to examine the distribution of pre-operative pain, function, and quality of life characteristics at time of TKR overall and stratified by baseline pain and function groups, with particular interest in the profiles of people with minimal pre-operative knee pain and high function; and 2) to determine whether the characteristics differ across groups.

MATERIAL AND METHODS

Population and National Registry

We used data from the Function and Outcomes Research for Comparative Effectiveness in TJR (FORCE-TJR), a national cohort of TJR patients from more than 130 surgeons in 22 sites nationwide. Parallel to the current US surgical practice, the majority of FORCE-TJR patients came from community-based surgeons, including fellowship-trained and general orthopedists in urban and rural locations, as well as teaching and non-teaching hospitals. Further description of the study population and patient-reported outcomes measures has been reported previously^{9,10}. In brief, the current study included all participants who had unilateral, primary TKRs with the surgical knee as the first surgical joint replaced between April 1, 2011 and May 31, 2014, and with osteoarthritis as the primary underlying reason for TKR.

Measures

Within 3 months before date of surgery, participants completed patient-reported outcomes (PRO) related to pre-operative pain, function and quality of life, as well as patient

demographics such as age, sex, race, education attainment, and medical status including body mass index (BMI) and comorbidities.

We assessed knee pain and function using the Knee Injury and Osteoarthritis Outcome Score (KOOS). KOOS is a self-administered knee-specific questionnaire containing 42 items that yields five separate subscales assessing domains of patient's pain (including frequency), symptoms (including stiffness, swelling, and clicking), activities of daily living for physical function, sport and recreation function, and knee-related quality of life^{11, 12}. For the overall KOOS, each item was scored from 0 to 4 and scores for all items were summed, transformed and normalized to a scale of 0 to 100, with 100 representing the best result (e.g., no pain; full function). A separate score was then calculated and normalized for each of the five subscales, with a scale of 0 to 100. For each subscale, if less than 50% of items were missing data in a subscale, then a valid subscale was calculated based on the developer's instructions¹³. If more than 50% of the items within a subscale were omitted, the response was considered invalid and no subscale score was calculated¹³. In addition, we derived participants' WOMAC scores based on their responses to the appropriate items from the KOOS.

We also assessed function using the Short-Form 36-item (SF-36) version 2. The SF-36 is a multi-purpose self-administered questionnaire with 36 items that yields eight subscales covering general health, physical function, social function, bodily pain, vitality, physical, mental, and emotional health, as well as two measures of summary scores for the physical (PCS) and mental (MCS) components^{14–16}. Using proprietary methods, each sub-scale and summary measure was scored and summed, transformed and normalized with a mean of 50 and one standard deviation (SD) of 10, reflecting that of the general population.. The SF-36 subscales and summary scores ranged from 0 to 100, with higher scores indicating better outcomes. Scores above 50 are better while scores below 50 are worse than the general population average^{17, 18}. Thus, a PCS function score of 40 indicates that function is 1 SD below the population average.

Method of Analysis

For the purpose of examining pre-operative patient symptom profiles, we used the KOOS pain subscale score and the SF-36 Physical Component Summary score (PCS). The SF 36 is a generic health questionnaire used extensively in populations with OA, especially those undergoing knee replacements¹⁹. The PCS and the MCS equivalence can also be derived from the shorter version SF12²⁰. Thus, we chose the PCS as a measure of physical function for possible comparisons across studies where SF 36 or SF 12 may be used. Further, the SF-36 PCS is a commonly used global measure of physical function; thus, we use PCS as the primary TKR function outcome for comparability. The SF-36 pain subscale, however, is not joint specific; thus, we use the KOOS pain subscale as our primary TKR pain outcome measure given it is specific to the surgical knee.

Moreover, there are currently no established gold-standards with regard to cut points for pain and function thresholds that would require TKR intervention²¹. While population-based normative scores for the SF-36 are well established^{17,18}, that of KOOS subscales are less so^{22, 23}. Thus, we applied the population-based cut-point of 40 for PCS, representing one

standard deviation below the population norm of average function. For KOOS, we took the mid-point of two pre-operative KOOS pain scores: that of American patients undergoing anterior cruciate ligament reconstruction²⁴, with a mean KOOS Pain score of 80; and Swedish patients undergoing arthroscopic partial meniscectomy²⁵, with a mean KOOS Pain score of 60. Our KOOS Pain score cut-point of 70 also approximates the mean pre-operative KOOS Pain score plus two standard deviations in Swedish patients awaiting total knee replacements¹². In addition to the normalized score of 100 in KOOS indicate no pain symptom¹³, in a correspondence between our team and the developer of KOOS (Ewa Roos, 03/17/15), a KOOS score of 70 to < 100 indicates minimal pain, 50 to < 70 moderate pain, 25 to < 50 severe pain, and 0 to < 25 extreme pain. For our main analysis, we classified patients as having high vs. little pain if KOOS Pain <70 vs. ≥70, and poor vs. high physical function (SF-36 PCS <40 vs. ≥40). We then classified patients into four groups: 1) little pain and high function, 2) little pain but poor function, 3) high pain but high function, and 4) high pain and poor function.

Descriptive statistics including patient demographic and clinical characteristics of the four groups were calculated and compared. We also compared items from each of the KOOS and SF-36 subscales to explore possible differences by the four groups that may add insights into reasons for TKR. For continuous variables, we calculated means and standard deviations, as well as the medians and inter-quartile range (IQR). We checked for normality using histograms and examined the mean, median, and standard deviation distributions, and used the Skewness/Kurtosis test for normality. For normal distributions, we compared means across the four groups and used analysis of variance (ANOVA) to determine statistical differences in the means, and for non-normal distributions, we used the Kruskal-Wallis rank test. For discrete variables, we calculated proportions and used chi-square tests to determine statistical differences in proportions among the four study groups. We used a two-tailed p value of < 0.05 as indication of statistical significance.

Given no established cut points with regard to pain and function as gold-standard indicators of TKR, we also performed sensitivity analyses to evaluate different cut-points, compare our group classification with previously published group classifications, and evaluate the stability of the pre-operative profiles across low and high volume surgical sites. For example, because there is no consensus on “cut-offs”, we performed analyses where the threshold of KOOS pain score was increased to < 80 vs. ≥80, as well as lowered to both <65 vs. ≥65 and < 60 vs. ≥60, for indication of high vs. little pain while keeping that of PCS function cut point the same. We also used SF36 Physical Performance (PF) subscale < 40 vs. ≥40 instead of the PCS < 40 vs. ≥40 to determine high vs. poor physical function, while keeping the KOOS pain cut point of 70 similar to the main analysis. The PF subscale is specific to physical function, and does not include pain or role function. Moreover, in order to compare our group classifications with those of other researchers, we examined the distributions of the main four pre-operative pain and function groups by WOMAC overall score of <39 vs. ≥39 as indication of inappropriate vs. appropriate use of TKRs^{26, 27}. Finally, we examined the distributions of the main four pre-operative pain and function groups by high and low volume surgical sites, where low volume sites were first defined as sites with less than 100 primary and unilateral TKR surgeries per year, then as sites with less than 150 surgeries per year²⁸. These analyses allowed us to compare our results to some of

those from prior research. All analyses were performed using STATA V13 (StataCorp, College Station, TX).

RESULTS

After applying our exclusion criteria, 7,428 patients remained eligible for the current study (Figure 1). Of these, 6,936 (93%) had non-missing data on both the pre-operative KOOS pain subscale and SF 36 PCS summary scores. General descriptive statistics for participants included in the study are as follows: the mean (SD) for age was 66.5 (9.2), BMI was 31.5 (6.1), females made up 62% of the study population, Caucasians made up 92% of the study population, 66% of the participants had some college education, and 28% used an assistive device such as a cane, walker, or wheelchair. The mean (SD) for KOOS pain subscore was 46.7 (17.5) and for SF-36 PCS was 33.0 (8.3). We also compared the above characteristics in the group of people eligible for the study (n=7,428) with the group of people actually included in the study with non-missing KOOS and PCS data (n=6,936). No clinically relevant differences in demographic factors and baseline medical characteristics were identified.

Demographic characteristics by groups of pre-operative pain and function are shown in Table 1. The majority (77%) of patients reported high pre-operative pain and poor physical function or the traditional indicators for TKR surgery (Group 4), 15% had high pain but high function (Group3), 4% had little pain but poor function (Group 2), and 5% had little pre-operative pain and high pre-operative function (Group 1). Compared with people in Group 4, the people in Group 1 on average were older, less obese, more likely men, and more highly educated. Descriptive statistics for those in Groups 2 and 3 were somewhere in between Groups 1 and 4.

Profiles of patients' medical status and general health are presented in Table 2. Compared with people with high pain and poor function in Group 4, people with little pain and high function in Group 1 were more likely to report being in excellent or very good health, and feeling happy or full of life most of the time, and were less likely to use an assisted device. Those in Group 1 were more similar to those with high pain but high function in Group 3 in responses to general health, and use of assisted device; those with little pain but poor function in Group 2 were more similar to those in Group 4 on these items. All groups scored low on the Charlson Co-morbidity index, and most differed from Group 4 in feeling full of life all or most of the time.

Pain and function profiles are presented in Table 3. Despite people in Group 1 having little pre-operative pain and high function, many in this group reported that they 'always' experienced pain or experienced it on a daily basis (48%), and a high proportion reported being 'constantly aware' of their knee problem or aware on a daily basis (86%). Despite this awareness, a very small percentage of this group experienced severe or extreme pain on stairs (5%) or pain in bed (1%). Moreover, over half felt that their health now limited them a lot in vigorous activities such as running or participating in strenuous sports (58%) and this was confirmed with severe or extreme difficulty on the KOOS sports items including

running (56%) and jumping (54%). Differences for all characteristics among the four groups were statistically significant at $P < 0.05$.

Results of the sensitivity analysis with the higher KOOS cut-off score (< 80 vs. ≥ 80) as indication of high vs. little pain, are presented in the Appendix Table 1. As expected, the proportion of people now considered as having little pre-operative pain and high pre-operative function decreased; more specifically, this group now made up only 2% of the study population.. As for using the KOOS Pain cut-point of 60, results are presented in Appendix Table 2. There were 10% of people in the group having little pre-operative pain and high pre-operative function; the inference with regard to overall general health and problems with pain frequency and vigorous activities remained the same as with using the cut-point of 70. A summary by the three cut points are presented in Appendix Table 3. Using the cut-point of 65 identified 7% of people in this group having little pre-operative pain and high pre-operative function but the inference remained the same (data not shown). The sensitivity analysis of using SF-36 PF subscale instead of the PCS summary score yielded almost identical results compared with the main analysis regarding proportion of TKR patients in each of the pre-operative pain and pre-operative function groups, as well as the profiles or distribution of characteristics within the groups (data not shown). We found that Hawker's definition²⁶ of inappropriate vs. appropriate TKR using the overall WOMAC score (in contrast to using the pain and function sub-scores in our main analysis) < 39 vs. ≥ 39 , resulted in 30% of our study population being classified as inappropriate TKR use. Only 5% of this group, however, overlapped with our group consisting of people with little pre-operative pain and high function. Finally, stratifying by low volume sites compared with high volume sites in our study population yielded similar results whether low volume sites were defined as < 100 primary and unilateral TKRs per year or < 150 surgeries per year: approximately 5% in the little pain and high function group (Group 1); 20% with either little pain or high function (Groups 2 and 3), and 75% in the high pain and poor function group (Group 4).

DISCUSSION

To our knowledge, this study is the first to examine patient-centered pre-operative symptom data with specific patient profiles at the time patients elected total knee replacements in a contemporary national US sample. In this analysis of a national sample of TKR patients from over 130 surgeons, the overwhelming majority of TKRs were performed among people reporting with both high pre-operative pain and poor pre-operative function. More than 75% of TKR patients reported both significant knee pain and limited function. An additional 20% of patients reported either significant pain or functional limitation. Of note, 5% of TKR utilization was among patients with neither significant pre-operative pain nor functional limitation. While this latter group in general did not experience pain on specific activities, a large proportion reported pain constantly or on a daily basis, and an overwhelming majority reported being always aware of their knee problem or aware on a daily basis. The people in this small group were generally older, more likely to be male, healthier, happier, and less disabled than those who had high pre-operative pain and low pre-operative function.

Results from our study add important insights to the growing debate about appropriate use of TKR. While deciding whether TKR is appropriate for a patient is based on clinical assessment as well as patient symptoms, our data highlight the patients' significant symptoms and profiles at the time they elected TKR. In contrast, using the modified-RAND definition of inappropriate surgical intervention criteria developed by Escobar⁶, Riddle et al^{7, 29} found using OAI data that 44% of TKRs were classified as appropriate, 22% were classified as inconclusive, and 34% as inappropriate use. However, in another recent study where Ghomrawi et al. also used Escobar's criteria²⁷, they found approximately 8% classified as inappropriate TKR use, 12% uncertain, and 81% as appropriate. Additionally, comparing the appropriateness criteria from Escobar and that from Hawker²⁶, Ghomrawi found only 5% of TKR utilization classified as inappropriate by both criteria. Even among experts, there is little agreement with regard to the definition of appropriate utilization of TKRs⁸.

In contrast to prior analyses using appropriateness criteria that combine clinical and radiographic information with patient symptom severity, we evaluate only the severity of patient-reported pain and functional limitations. Despite this difference, our patient-centered analyses have interesting parallels with the results using standard appropriateness tools. Our analyses of patient symptom profiles are most similar to that of Ghomrawi's results where two appropriateness criteria overlapped. We found that 5% of our patients selected for TKRs had little pain and high function using KOOS of 70 and SF 36 PCS of 40 cut points, and that all in this 5% subgroup overlapped with Hawker's criteria of TKR inappropriateness using the overall WOMAC score. In contrast to Hawker's method of using aggregate scores²⁶, we evaluated the pre-operative profiles using pain and function sub-scores independently. While the majority of patients reported both severe pain and disability, we found distinct sub-groups of patients who had either severe pain or functional limitation, but not both. Of note, over 50% of study sites in the FORCE TJR cohort were from community-based surgeons in 22 states; thus, FORCE TJR may be more representative of the contemporary US surgical practices than populations in Riddle and Ghomrawi's studies. Still, when we stratified our analyses by low volume and high volume surgical sites, our results were very similar to the main analysis, suggesting that the patients' criteria were consistent across surgical practice. It is possible that patients who had the traditional indicators of high pain and poor function for TKRs had surgical interventions for the purpose of relieving pain and restoring physical function, and a small proportion of people had TKRs for the purpose of improving quality of life and reclaiming the active lifestyle that they previously had, despite having little pre-operative pain and high pre-operative function.

Several factors may impact differences in defining appropriateness of TKR utilization. First, Riddle's definitions of slight and moderate symptoms were based on quartile splits of the combined WOMAC pain and function scales. These quartile splits, while applicable to Riddle's study population, may be different in other study populations, especially if there are cultural, age, or sex differences, as well as differences in clinically relevant thresholds for pain and function, as suggested by Ghomrawi²⁷. It is also possible that different cultures may have different thresholds for what is considered severe pain or limited physical function, and these thresholds may not be captured appropriately using clinical tools currently available. Nevertheless, pain and functional limitations may not be the only

indicators for TKRs, especially if the pain is not due to knee OA. Moreover, severity of OA on xray does not correlate with symptom severity so cannot add to this determination^{30, 31}. In addition, Riddle's heavy weighting of pre-surgical radiographic grades in his definition of appropriateness may not be widely applicable to other clinicians in determining appropriate timing of TKRs for their patients. It is also possible that the use of appropriateness criteria that were based on evidence published prior to 1999 may not be applicable to current clinical practice since today's US patients are significantly younger and more obese⁸. Ultimately, a sub-group of patients and clinicians may consider loss of quality of life and for general health maintenance as the primary indication for joint replacement. Indeed, once patients indicate that they cannot enjoy activities of daily living or any recreational activities, certain clinicians would consider that as appropriate reason for TKRs³².

Furthermore, the optimal time for surgical interventions such as TKR is still widely debated, and remains a decision best made between the clinicians and their patients. Some researchers have warned about the negative impact that delaying surgical interventions may have on functional gain and pain relief after TKRs, as those with poorest function before surgery were more likely to have smaller functional gains^{33,34}. Surgeons and patients must balance delays in surgery timing with the fact that patients had lower pain relief and smaller functional improvement than were patients considered appropriate for TKRs^{8, 35}. It is worth noting that using the modified Escobar's appropriateness criteria, Riddle et al.²⁹ recently found that the group classified as having inconclusive use of TKR had very similar pre-operative scores as the group classified as having appropriate TKR use on WOMAC physical function, KOOS symptoms, and KOOS pain scales..

Several strengths and limitations of our current study warrant discussion. First, there is no standard measure of pain or functional limitations in patient-reported outcomes. We selected the SF PCS because it is a commonly used measure of physical function and for its well-known population-based normative scores^{17,18}. Use of the PCS allows us to compare patients in our cohorts to prior TKR research. While a PCS score below 50 indicates that function is below the general population average, there is no gold standard regarding the threshold for TKR utilization. Thus, we used a PCS cut-point of 40 because it is 1 SD below the population-based average of 50. Also, a PCS cut-point of 40 is also approximately 1 SD below the mean FORCE cohort pre-op PCS (mean=33 + 8.3 SD). As for a measure of pain, we selected the KOOS pain subscale for its joint-specific properties. Since the population-based normative score for KOOS pain score is less well established, we first applied a KOOS pain cut point of 70 which was 2 SD above the mean KOOS pain score among Swedish patients awaiting TKR¹². We then applied several cut points in our sensitivity analyses to ascertain the robustness of our analyses, including KOOS pain score cut-point of 80, 65, and 60. Despite lowering the cut-point for KOOS pain score to that of 60, inference from the results was similar to that of using 70 as a cut point. Second, we had no radiograph findings to determine the biologic severity of the diseased knee in our assessment; however, all participants selected for our current study reported underlying OA as the reason for TKRs. While radiograph documentation of OA changes is important, imaging severity classifications do not correlate strongly with patient-reported pain and disability, and may not be a consistent predictor of TKR use^{30, 31}. Finally, pain and functional limitations and radiographic findings may not be the only indicators of TKRs. Thus, the decision for

surgical interventions is based on a dialogue between doctor and patient about patient-specific knee OA symptoms, goals, and expectations. Identifying additional clinical reasons for TKR, such as mal-alignment or other deformities, requires clinical chart reviews or review of radiographs, which were not available in this study. Also, appropriateness of TKR use may be related to expected gains in post-op physical function and pain relief, which were not examined in the current analyses. Nevertheless, we evaluated pre-operative pain and pre-operative function using standardized methods of assessments in a nationally representative sample of thousands of patients who underwent primary and unilateral TKRs. It is important to note that while FORCE-TJR patients completed the surveys before surgery, the scores were not available to the surgeons or their patients at the time that they made their decision regarding surgical intervention³².

CONCLUSIONS

Our study described important characteristics of patients selected for knee arthroplasty in a contemporary and nationally-representative cohort. While the vast majority of TKR patients reported high pain and/or poor function before surgery, a small minority did not. This small group of people were more likely to be older, less obese, more males, better educated, healthier, happier, and less disabled than those with high pain and poor function before surgery; yet, they were constantly aware of their knee problems and were always experiencing pain on a daily basis. Quality of life issues and limitation on vigorous activities could be possible reasons why the 5% with low pre-operative pain and high pre-operative function elected TKRs. Further investigation is needed to determine the reasons for TKR use, especially among patients with little pain and high function prior to surgery, as this may have important policy implications. Updated criteria for appropriateness of TKRs that can be used by patients and surgeons prior to the decision to operate can support shared decision-making and assure optimal timing and use of TKR for US patients suffering from advanced knee OA.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We gratefully acknowledge the more than 130 surgeons who participate in the FORCE-TJR national consortium, especially the core site Principal Investigators, Courtland Lewis MD, Philip Noble PhD, Regis O'Keefe MD, PhD, and Vincent Pellegrini MD. We also appreciate the editorial assistance of Sylvie Puig PhD, and statistical assistance of Ning Guo, MS.

Grants: This study is supported by a grant from the Agency for Healthcare Research and Quality (AHRQ) P50HS018910 and U-S Nguyen is partly funded by the Arthritis Foundation and NIAMS K01AR064352. The funding sources had no involvement in the: study design; data collection, analysis, or interpretation; writing of the report; or submission of the article for publication.

Appendix

Appendix Table 1

Characteristics by Groups—Using KOOS <80 vs. ≥80 and PCS <40 vs. ≥40 to define pain and function status.

Characteristics*	Group 1 Little Pain High Function n=141 (2%)	Group 2 Little Pain Poor Function n=98 (1%)	Group 3 High Pain High Function n= 1236 (18%)	Group 4 High Pain Poor Function n=5461 (79%)
Demographics				
Age (Years, Mean (SD))	71.0 (7.3)	71.3 (7.9)	66.8 (6.4)	66.2 (9.3)
BMI (Mean (SD))	28.9 (5.4)	30.9 (5.7)	29.6 (5.3)	32.0 (6.2)
Sex (Female)	41%	60%	49%	65%
Race (White)	94%	89%	94%	91%
Post-High School Education	70%	64%	73%	65%
Insurance Type (Medicare)	72%	67%	52%	54%
General Health and Mental Status				
General Health SF36	56.5 (6.2)	49.8 (8.7)	55.0 (6.11)	48.7 (8.9)
Excellent/Very Good	72%	50%	70%	40%
Co-Morbidity Index Median (IQR)	0 (0,1)	1 (0,2)	0 (0,1)	0 (0,1)
Disability (use of assisted device)	9%	40%	10%	32%
Mental Health, SF36MCS, Mean(SD)	55.7 (7.5)	53.7 (7.5)	52.4 (9.3)	49.0 (10.6)
Feel Depressed-All/Most of the time	1%	2%	3%	6%
Feel Happy-All/Most of the time	89%	76%	77%	65%
Vitality, SF36 VT, Mean (SD)	57.8 (8.5)	49.1 (9.4)	53.6 (8.7)	44.7 (10.0)
Feel Full of Life-All/Most of the time	76%	59%	68%	39%
Pain and Symptoms				
Pain Frequency Daily/Always	27%	29%	89%	97%
Awareness Knee Problem Daily/Constantly	77%	88%	97%	99%
Pain, KOOS Mean (SD)	87.7 (6.0)	87.0 (6.7)	56.6 (13.4)	42.7 (15.4)
Pain on stairs (severe/extreme)	0	3%	29%	67%
Pain in bed (severe/extreme)	0	0	10%	21%
Symptoms, KOOS Mean (SD)	80.6 (11.3)	78.5 (10.3)	62.3 (14.6)	52.8 (14.8)
Stiffness Late in Day(severe/extreme)	3%	2%	22%	46%
Swelling of the Knee (severe/extreme)	18%	16%	34%	48%
Clicking/Grinding Knee (severe/extreme)	16%	19%	48%	63%
Function				
Function, SF36 PCS, Mean (SD)	47.9 (4.8)	33.5 (4.7)	44.7 (3.7)	30.0 (6.0)
Vigorous Activities - Limited a lot	55%	88%	59%	89%
Moderate Activities - Limited a lot	1%	26%	3%	38%
Walking > 1 mile - Limited a lot	28%	76%	29%	83%
Function, KOOS-Sports, Mean (SD)	50.5 (22.8)	28.1 (21.4)	29.4 (19.0)	14.3 (16.2)
Running - Severe/Extreme Difficulty	46%	81%	77%	93%
Jumping - Severe/Extreme Difficulty	45%	80%	76%	93%

* All P values were < 0.05, except race

Appendix Table 2

Characteristics by Groups—Using KOOS <60 vs. ≥60 and PCS <40 vs. ≥40 to define pain and function status.

Characteristics*	Group 1 Little Pain High Function n=702 (10%)	Group 2 Little Pain Poor Function n=828 (12%)	Group 3 High Pain High Function n= 675 (10%)	Group 4 High Pain Poor Function n=4731 (68%)
Demographics				
Age (Years, Mean (SD))	68.6 (7.9)	69.9 (8.6)	65.8 (8.7)	65.7 (9.3)
BMI (Mean (SD))	29.4 (5.1)	30.9 (5.8)	29.9 (5.5)	32.2 (6.2)
Sex (Female)	45%	55%	52%	67%
Race (White)	94%	95%	93%	90%
Post-High School Education	76%	70%	70%	64%
Insurance Type (Medicare)	59%	64%	49%	52%
General Health and Mental Status				
General Health SF36	55.7 (6.0)	50.6 (8.0)	54.7 (6.3)	48.4 (9.1)
Excellent/Very Good	72%	47%	68%	39%
Co-Morbidity Index Median (IQR)	0 (0,1)	0 (0,1)	0 (0,1)	0 (0,1)
Disability (use of assisted device)	9%	29%	10%	33%
Mental Health, SF36MCS, Mean(SD)	56.5 (8.5)	56.8 (9.4)	52.3 (11.1)	50.7 (12.5)
Feel Depressed-All/Most of the time	2%	2%	4%	7%
Feel Happy-All/Most of the time	82%	78%	72%	62%
Vitality, SF36 VT, Mean (SD)	55.7 (8.0)	49.1 (8.9)	52.2 (9.2)	44.0 (10.0)
Feel Full of Life-All/Most of the time	73%	55%	64%	37%
Pain and Symptoms				
Pain Frequency Daily/Always	69%	78%	96%	98%
Awareness Knee Problem Daily/Constantly	92%	96%	98%	99%
Pain, KOOS Mean (SD)	72.2 (9.5)	69.2 (8.3)	46.9 (9.9)	39.0 (12.9)
Pain on stairs (severe/extreme)	9%	21%	45%	74%
Pain in bed (severe/extreme)	2%	1%	16%	24%
Symptoms, KOOS Mean (SD)	71.5 (13.2)	69.0 (12.8)	56.5 (13.6)	50.5 (13.7)
Stiffness Late in Day(severe/extreme)	7%	11%	34%	52%
Swelling of the Knee (severe/extreme)	23%	26%	43%	51%
Clicking/Grinding Knee (severe/extreme)	34%	40%	57%	66%
Function				
Function, SF36 PCS, Mean (SD)	45.9 (4.3)	33.1 (5.1)	44.2 (3.4)	29.5 (6.0)
Vigorous Activities - Limited a lot	60%	88%	57%	89%
Moderate Activities - Limited a lot	3%	24%	3%	40%
Walking > 1 mile - Limited a lot	28%	78%	29%	83%
Function, KOOS-Sports, Mean (SD)	39.8 (20.4)	26.0 (19.1)	23.0 (16.6)	12.5 (16.0)
Running - Severe/Extreme Difficulty	63%	84%	85%	95%
Jumping - Severe/Extreme Difficulty	62%	84%	85%	94%

* All P values were < 0.05, except race

Appendix Table 3

Summary by the three different KOOS Pain Cut-points

Characteristics	Pain of 80 PCS of 40	Pain of 70 PCS of 40	Pain of 60 PCS of 40
Little Pain and High Function Group (of total n)	2%	5%	10%
Within the Little Pain and High Function Group			
General Health – Very Good/Excellent	72%	73%	72%
Pain frequency - Daily/Always	27%	48%	69%
Awareness of knee problems - Daily/Constantly	77%	86%	92%
Vigorous activity - Limited a lot	55%	58%	60%
Running - Severe/Extreme Difficulty	46%	56%	63%
Jumping - Severe/Extreme Difficulty	45%	54%	65%

REFERENCES

1. NIH Consensus Statement on total knee replacement. NIH Consens State Sci Statements. 2003; 20(1):1–34. [PubMed: 17308549]
2. Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). Rate of Total Knee Replacement for Persons Aged >65 Years, by Sex --- United States, 1979--2002. MMWR. 2005; 54(07):179.
3. Kim S. Changes in surgical loads and economic burden of hip and knee replacements in the US: 1997–2004. Arthritis & Rheumatism. 2008; 59(4):481–488. [PubMed: 18383407]
4. Racial disparities in total knee replacement among Medicare enrollees--United States, 2000–2006. MMWR Morb Mortal Wkly Rep. 2009; 58(6):133–138. [PubMed: 19229164]
5. Number of all-listed procedures for discharges from short-stay hospitals, by procedure category and age: United States, 2010. 2010.
6. Escobar A, Quintana JM, Arostegui I, Azkarate J, Guenaga JI, Arenaza JC, et al. Development of explicit criteria for total knee replacement. Int J Technol Assess Health Care. 2003; 19(1):57–70. [PubMed: 12701939]
7. Riddle DL, Jiranek WA, Hayes CW. Use of a validated algorithm to judge the appropriateness of total knee arthroplasty in the United States: a multicenter longitudinal cohort study. Arthritis Rheumatol. 2014; 66(8):2134–2143. [PubMed: 24974958]
8. Katz JN. Editorial: appropriateness of total knee arthroplasty. Arthritis Rheumatol. 2014; 66(8): 1979–1981. [PubMed: 24975039]
9. Franklin PD, Allison JJ, Ayers DC. Beyond joint implant registries: a patient-centered research consortium for comparative effectiveness in total joint replacement. JAMA. 2012; 308(12):1217–1218. [PubMed: 23011710]
10. Franklin PD, Lewallen D, Bozic K, Hallstrom B, Jiranek W, Ayers DC. Implementation of patient-reported outcome measures in u.s. Total joint replacement registries: rationale, status, and plans. J Bone Joint Surg Am. 2014; 96(Suppl 1):104–109. [PubMed: 25520425]
11. Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. Health Qual Life Outcomes. 2003; 1:64. [PubMed: 14613558]
12. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes. 2003; 1:17. [PubMed: 12801417]
13. The 2012 User's Guide to: Knee injury and Osteoarthritis Outcome Score KOOS. <http://www.koos.nu/KOOSusersguide2012.pdf>

14. Sabry AA, Elbasyouni SR, Kalil AM, Abdel-Rahim M, Mohsen T, Sleem A. Markers of inflammation and atherosclerosis in Egyptian patients with systemic lupus erythematosus. *Nephrology*. 2006; 11(4):329–335. [PubMed: 16889573]
15. Haley SM, McHorney CA, Ware JE Jr. Evaluation of the MOS SF-36 physical functioning scale (PF-10): I. Unidimensionality and reproducibility of the Rasch item scale. *J Clin Epidemiol*. 1994; 47(6):671–684. [PubMed: 7722580]
16. McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care*. 1994; 32(1):40–66. [PubMed: 8277801]
17. Ware, JE., Jr; Kosinski, M. SF-36 Physical & Mental Health Summary Scales: A Manual for Users of Version 1. 2nd. Lincoln, RI: QualityMetric; 2001.
18. Ware JE Jr. SF-36 Health Survey Update. *SPINE*. 2000; 25:3130–3139. [PubMed: 11124729]
19. Desmeules F, Dionne CE, Belzile E, Bourbonnais R, Frémont P. Waiting for total knee replacement surgery: factors associated with pain, stiffness, function and quality of life. *BMC Musculoskeletal Disorders*. 2009; 10:52. [PubMed: 19457252]
20. Ware JE Jr, Kosinski M, Keller SD. A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care*. 1996; 34:220–233. [PubMed: 8628042]
21. Gossec L, Paternotte S, Maillefert JF, Combescure C, Conaghan PG, Davis AM, et al. The role of pain and functional impairment in the decision to recommend total joint replacement in hip and knee osteoarthritis: an international cross-sectional study of 1909 patients. Report of the OARSI-OMERACT Task Force on total joint replacement. *Osteoarthritis Cartilage*. 2011; 19(2):147–154. [PubMed: 21044689]
22. Cameron KL, Thompson BS, Peck KY, Owens BD, Marshall SW, Svoboda SJ. Normative values for the KOOS and WOMAC in a young athletic population: history of knee ligament injury is associated with lower scores. *Am J Sports Med*. 2013; 41(3):582–589. [PubMed: 23328737]
23. Paradowski PT, Bergman S, Sunden-Lundius A, Lohmander LS, Roos EM. Knee complaints vary with age and gender in the adult population. Population-based reference data for the Knee injury and Osteoarthritis Outcome Score (KOOS). *BMC Musculoskeletal Disord*. 2006; 7:38. [PubMed: 16670005]
24. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)--development of a self-administered outcome measure. *J Orthop Sports Phys Ther*. 1998; 28(2):88–96. [PubMed: 9699158]
25. Roos EM, Roos HP, Ekdahl C, Lohmander LS. Knee injury and Osteoarthritis Outcome Score (KOOS)--validation of a Swedish version. *Scand J Med Sci Sports*. 1998; 8(6):439–448. [PubMed: 9863983]
26. Hawker GA, Wright JG, Coyte PC, Williams JI, Harvey B, Glazier R, et al. Determining the need for hip and knee arthroplasty: the role of clinical severity and patients' preferences. *Med Care*. 2001; 39(3):206–216. [PubMed: 11242316]
27. Ghomrawi HM, Alexiades M, Pavlov H, Nam D, Endo Y, Mandl LA, et al. Evaluation of two appropriateness criteria for total knee replacement. *Arthritis Care Res (Hoboken)*. 2014; 66(11): 1749–1753. [PubMed: 24964968]
28. Riddle DL, Hayes CW, Jiranek WA. Response to letter by Franklin and colleagues. *Arthritis Rheumatol*. 2014
29. Riddle DL, Perera RA, Jiranek WA, Dumenci L. Using surgical appropriateness criteria to examine outcomes of total knee arthroplasty in a United States sample. *Arthritis Care Res (Hoboken)*. 2015; 67(3):349–357. [PubMed: 25132662]
30. Bedson J, Croft PR. The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. *BMC Musculoskeletal Disord*. 2008; 9:116. [PubMed: 18764949]
31. Finan PH, Buenaver LF, Bounds SC, Hussain S, Park RJ, Haque UJ, et al. Discordance between pain and radiographic severity in knee osteoarthritis: findings from quantitative sensory testing of central sensitization. *Arthritis Rheum*. 2013; 65(2):363–372. [PubMed: 22961435]
32. Fischer, K. One third of knee replacements for arthritis may be unnecessary. 2014. <http://www.healthline.com/health-news/one-third-knee-replacements-unnecessary-070114#1>

33. Fortin PR, Penrod JR, Clarke AE, St-Pierre Y, Joseph L, Belisle P, et al. Timing of total joint replacement affects clinical outcomes among patients with osteoarthritis of the hip or knee. *Arthritis Rheum.* 2002; 46(12):3327–3330. [PubMed: 12483739]
34. Jones CA, Voaklander DC, Suarez-Alma ME. Determinants of function after total knee arthroplasty. *Phys Ther.* 2003; 83(8):696–706. [PubMed: 12882610]
35. Franklin P, Nguyen U, Ayers D, Weissman N. Letter in response to the arthritis and rheumatology paper (Riddle, Jiranek et al) and Katz commentary; Appropriateness of TKR utilization. *Arthritis Rheumatol.* 2014 Nov 4.

Figure 1.
Flow-Chart of FORCE-TJR Study Population

Table 1

Demographic Characteristics by Pre-Operative Pain and Function Status Using KOOS <70 vs. ≥70 and PCS <40 vs. ≥40 to define pain and function status.

Characteristics*	Group 1 Little Pain High Function n=327 (5%)	Group 2 Little Pain Poor Function n=250 (4%)	Group 3 High Pain High Function n=1050 (15%)	Group 4 High Pain Poor Function n=5309 (77%)
Demographics				
Age (Years, Mean (SD))	69.6 (7.7)	70.3 (8.6)	66.5 (8.5)	66.1 (9.3)
BMI (Mean (SD))	29.0 (5.2)	30.7 (5.7)	29.7 (5.4)	32.0 (6.2)
Sex (Female)	45%	54%	50%	66%
Race (White)	94%	92%	94%	91%
Post-High School Education	75%	69%	72%	65%
Insurance Type (Medicare)	63%	63%	52%	54%

* All P values were < 0.05, except race

Table 2

Medical Status and General Health Profiles by Pre-Operative Pain and Function Status Using KOOS <70 vs. >=70 and PCS <40 vs. >= 40 to define pain and function status.

Characteristics*	Group 1 Little Pain High Function n=327 (5%)	Group 2 Little Pain Poor Function n=250 (4%)	Group 3 High Pain High Function n=1050 (15%)	Group 4 High Pain Poor Function n=5309 (77%)
General Health and Mental Status				
General Health, SF36 GH, Mean (SD)	56.1 (6.2)	50.9 (8.4)	54.9 (6.1)	48.6 (9.0)
Excellent/Very Good	73%	48%	69%	40%
Co-Morbidity Index Median (IQR)	0 (0, 1)	0 (0, 2)	0 (0, 1)	0 (0, 1)
Disability (use assisted device)	9%	32%	10%	33%
Mental Health, SF36MCS, Mean (SD)	56.5 (9.3)	57.0 (9.7)	53.8 (10.3)	51.3 (12.3)
Feel Depressed-All/Most of time	3%	2%	3%	6%
Feel Happy-All/Most of time	80%	77%	76%	64%
Vitality, SF36 VT, Mean (SD)	56.8 (8.0)	50.2 (9.0)	53.1 (8.8)	44.5 (9.9)
Feel Full of Life-All/Most of time	75%	60%	67%	39%

* All P values were < 0.05

Table 3

Pain and Function Status Profiles by Pre-Operative Pain and Function Status Using KOOS <70 vs. ≥70 and PCS <40 vs. ≥40 to define pain and function status.

Characteristics*	Group 1 Little Pain High Function n=327 (5%)	Group 2 Little Pain Poor Function n=250 (4%)	Group 3 High Pain High Function n=1050 (15%)	Group 4 High Pain Poor Function n=5309 (77%)
Pain and Symptoms				
Pain Frequency Daily/Always	48%	54%	93%	97%
Awareness Knee Problem				
Daily/Constantly	86%	92%	98%	99%
Pain, KOOS Mean (SD)	80.3 (7.8)	79.4 (7.6)	53.4 (11.9)	41.8 (14.6)
Pain on stairs (severe/extreme)	5%	8%	33%	69%
Pain in bed (severe/extreme)	1%	1%	11%	22%
Symptoms, KOOS Mean (SD)	76.2 (12.5)	74.6 (11.6)	60.4 (14.2)	52.2 (14.5)
Stiffness Late in Day(severe/extreme)	4%	7%	25%	47%
Swelling (severe/extreme)	21%	20%	36%	49%
Clicking/Grinding (severe/extreme)	26%	29%	51%	64%
Function				
Function, SF36 PCS, Mean (SD)	47.1 (4.7)	33.8 (4.6)	44.4 (3.5)	29.8 (6.0)
Vigorous Activities - Limited a lot	58%	86%	59%	89%
Moderate Activities - Limited a lot	3%	25%	3%	38%
Walking > 1 mile - Limited a lot	29%	77%	29%	83%
Function, KOOS-Sports, Mean (SD)	44.1 (22.0)	28.4 (21.1)	27.7 (18.3)	13.9 (15.9)
Running - Severe/Extreme Difficulty	56%	80%	80%	94%
Jumping - Severe/Extreme Difficulty	54%	79%	79%	93%

* All P values were < 0.05