

# UCSF

## UC San Francisco Previously Published Works

### Title

Examining Timeliness of Total Knee Replacement Among Patients with Knee Osteoarthritis in the U.S.: Results from the OAI and MOST Longitudinal Cohorts.

### Permalink

<https://escholarship.org/uc/item/13j578xg>

### Journal

Journal of Bone and Joint Surgery, 102(6)

### ISSN

0021-9355

### Authors

Ghomrawi, HMK  
Mushlin, AI  
Kang, R  
[et al.](#)

### Publication Date

2020-03-18

### DOI

10.2106/jbjs.19.00432

Peer reviewed



A commentary by Michael G. Zywiol, MD, MSc, FRCSC, is linked to the online version of this article at [jbsj.org](http://jbsj.org).

# Examining Timeliness of Total Knee Replacement Among Patients with Knee Osteoarthritis in the U.S.

## Results from the OAI and MOST Longitudinal Cohorts

H.M.K. Ghomrawi, PhD, MPH, A.I. Mushlin, MD, ScM, R. Kang, MA, S. Banerjee, PhD, J.A. Singh, MD, L. Sharma, MD, C. Flink, MD, M. Nevitt, PhD, T. Neogi, MD, PhD,\* and D.L. Riddle, PT, PhD\*

*Investigation performed at Feinberg School of Medicine of Northwestern University, Chicago, Illinois*

**Background:** Patients with knee osteoarthritis may undergo total knee replacement too early or may delay or underuse this procedure. We quantified these categories of total knee replacement utilization in 2 cohorts of participants with knee osteoarthritis and investigated factors associated with each category.

**Methods:** Data were pooled from 2 multicenter cohort studies that collected demographic, patient-reported, radiographic, clinical examination, and total knee replacement utilization information longitudinally on 8,002 participants who had or were at risk for knee osteoarthritis and were followed for up to 8 years. Validated total knee replacement appropriateness criteria were longitudinally applied to classify participants as either potentially appropriate or likely inappropriate for total knee replacement. Participants were further classified on the basis of total knee replacement utilization into 3 categories: timely (indicating that the patient had total knee replacement within 2 years after the procedure had become potentially appropriate), potentially appropriate but knee not replaced (indicating that the knee had remained unreplaced for >2 years after the procedure had become potentially appropriate), and premature (indicating that the procedure was likely inappropriate but had been performed). Utilization rates were calculated, and factors associated with each category were identified.

**Results:** Among 8,002 participants, 3,417 knees fulfilled our inclusion and exclusion criteria and were classified into 1 of 3 utilization categories as follows: 290 knees (8% of the total and 9% of the knees for which replacement was potentially appropriate) were classified as “timely”, 2,833 knees (83% of the total and 91% of those for which replacement was potentially appropriate) were classified as “potentially appropriate but not replaced”, and 294 knees (comprising 9% of the total and 26% of the 1,114 total knee replacements performed) were considered to be “likely inappropriate” yet underwent total knee replacement and were classified as “premature”. Of the knees that were potentially appropriate but were not replaced, 1,204 (42.5%) had severe symptoms. Compared with the patients who underwent timely total knee

*continued*

\*T. Neogi, MD, PhD, and D.L. Riddle, PT, PhD, contributed equally as senior authors.

**Disclosure:** This project is funded by a grant (R21-AR069867) to the first author from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) and by a P-30 (P30-AR072579) grant from the National Institute of Arthritis and Musculoskeletal and Skin Diseases. The Osteoarthritis Initiative (OAI) is a public-private partnership comprised of 5 contracts (N01-AR-2-2258, N01-AR-2-2259, N01-AR-2-2260, N01-AR-2-2261, N01-AR-2-2262) funded by the National Institutes of Health (NIH) and conducted by the OAI Study Investigators. Private funding partners include Merck Research Laboratories, Novartis Pharmaceuticals Corporation, GlaxoSmithKline, and Pfizer, Inc. Private sector funding for the OAI is managed by the Foundation for the NIH. The Multicenter Osteoarthritis study (MOST) is comprised of 4 contracts (U01 AG18947, AG188832, AG19069, AG18820) funded by the NIH and conducted by the MOST Study Investigators. On the **Disclosure of Potential Conflicts of Interest** forms, which are provided with the online version of the article, one or more of the authors checked “yes” to indicate that the author had a relevant financial relationship in the biomedical arena outside the submitted work and “yes” to indicate that the author had other relationships or activities that could be perceived to influence, or have the potential to influence, what was written in this work (<http://links.lww.com/JBJS/F667>).

Disclaimer: This manuscript was prepared using an OAI public use data set (in addition to data obtained within NIH/NIAMS funded ancillary grants) and does not necessarily reflect the opinions or views of the OAI investigators, the NIH, or the private funding partners. This manuscript was prepared using MOST data and does not necessarily reflect the opinions or views of MOST investigators.

replacement, the likelihood of being classified as potentially appropriate but not undergoing total knee replacement was greater for black participants and the likelihood of having premature total knee replacement was lower among participants with a body mass index of  $>25 \text{ kg/m}^2$  and those with depression.

**Conclusions:** In 2 multicenter cohorts of patients with knee osteoarthritis, we observed substantial numbers of patients who had premature total knee replacement as well as of patients for whom total knee replacement was potentially appropriate but had not been performed  $>2$  years after it had become potentially appropriate. Further understanding of these observations is needed, especially among the latter group.

**Clinical Relevance:** Undergoing total knee replacement too early may result in little or no benefit while exposing the patient to the risks of a major operation, whereas waiting too long may cause limitations in physical activity that in turn increase the risk of additional disability and chronic disease; however, little is known about timing of this surgery. We quantified the extent of premature, timely, and delayed use, and found a high prevalence of both premature and delayed use.

Total knee replacement is a costly but effective elective surgical procedure<sup>1</sup>. Nearly 1,000,000 procedures are performed in the United States (U.S.) each year, and projections of a rapid increase by 2030 have concerned policy makers, with several initiatives aimed at restraining costs<sup>2-4</sup>.

On the other hand, total knee replacement underuse is often acknowledged as a high priority by policy makers<sup>5,6</sup>. Health insurance expansion under the Affordable Care Act is improving coverage for disadvantaged populations<sup>4,7</sup>; however, given that total knee replacement is a preference-based procedure, insurance

**TABLE I Elements of the Modified Escobar Appropriateness Criteria for Total Knee Replacement\***

Factor	Description
Age	
Level 1	<55 yr
Level 2	55 to 65 yr
Level 3	>65 yr
Knee stability	
Level 1	Preserved mobility and stable joint (<5° flexion contracture and normal or minor medial or lateral gapping in the 20° flexed knee)
Level 2	Limited mobility and/or unstable joint (≥5° flexion contracture and/or moderate or severe medial or lateral gapping in the 20° flexed knee)
Compartments involved	
Level 1	Unicompartmental
Level 2	Bicompartmental or tricompartmental
Radiographic findings	
Level 1	KL grade, <3
Level 2	KL grade, 3
Level 3	KL grade, 4
Symptomatology†	
Level 1	Slight (mild overall functional loss and function-related pain [e.g., up to half of WOMAC Pain and Physical Function scale items scored as mild and combined scale scores from 0 to 11])
Level 2	Moderate (moderate overall functional loss and function-related pain [e.g., up to half of WOMAC Pain and Physical Function scale items marked as moderate and combined scale scores from 12 to 22])
Level 3	Intense (intense overall functional loss and function-related pain [e.g., up to half of WOMAC Pain and Physical Function scale items marked as intense and combined scale scores from 23 to 33])
Level 4	Severe (severe overall functional loss and function-related pain [e.g., up to half of WOMAC Pain and Physical Function scale items marked as severe and combined scale scores ≥34])

\*Sixteen combinations of factors, depending on levels involved, determined whether the person was appropriate or inappropriate for surgery. For example, total knee replacement is appropriate for a ≥55-year-old patient with a KL (Kellgren-Lawrence) grade of 4 and intense or severe symptoms but it is inappropriate for a patient with KL grade of ≤3 and slight symptoms. †Symptomatology is based on the combined WOMAC Pain and Physical Function raw scores (score range, 0 to 88).

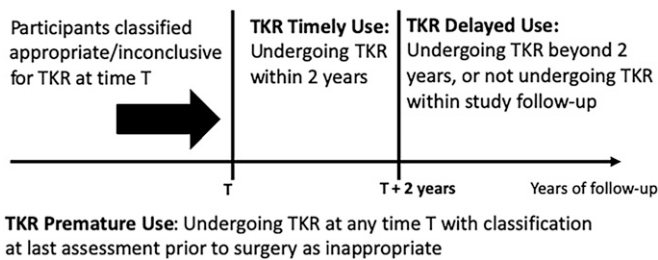


Fig. 1

Diagram showing the definitions of total knee replacement (TKR) utilization categories: timely, premature, and potentially appropriate but not replaced.

coverage alone will not fully address utilization issues related to additional barriers to treatment access and uptake<sup>8-12</sup>. Some patients may decide after careful consideration not to have total knee replacement despite severe knee pain, whereas others may elect total knee replacement in the early stage of disease in order to maintain functions that they consider to be critical.

Undergoing total knee replacement too early may result in little or no benefit while exposing the patient to the risks of a major operation, whereas waiting too long may cause limitations in physical activity that in turn increase the risk of additional disability and chronic disease<sup>13,14</sup>. Appropriate engagement in a preference-based and evidence-based decision that includes a full accounting of the risks and benefits of both pathways (total knee replacement and no total knee replacement) should minimize these adverse consequences.

Few reports have assessed the appropriateness of total knee replacement among those who have undergone the procedure, or have quantified the premature use of total knee replacement<sup>15-17</sup>, and no studies, to our knowledge, have quantified delayed use or underuse. Available information in this area is based principally on assumptions from studies examining racial, ethnic, and regional total knee replacement disparities in the rates of actual procedures<sup>7,12,15</sup> but not in relation to potentially eligible populations. For example, it is unclear if the utilization gap between whites and blacks is due to a higher proportion of eligible blacks delaying or underusing total knee replacement, more whites prematurely using total knee replacement, or a mix of both. In addition, it is unclear if factors that have been associated with disparities in the utilization of total knee replacement also affect the timeliness of total knee replacement or should be targets for intervention.

We aimed to quantify the appropriateness and timeliness of total knee replacement in 2 large multicenter cohorts of individuals with or at high risk for developing knee osteoarthritis and to investigate the associated predictors.

## Materials and Methods

### Data Sources

We pooled data from 2 cohort studies of individuals with or at high risk of developing knee osteoarthritis that involved similar data-collection protocols. The Osteoarthritis Initiative (OAI) is a prospective longitudinal cohort study of persons recruited from the community who either had

symptomatic and radiographic knee osteoarthritis or were at high risk for developing symptomatic and radiographic knee osteoarthritis. The OAI study enrolled 4,976 persons ranging from 45 to 79 years of age from February 2004 to May 2006 at 4 clinical sites (Baltimore, MD; Columbus, OH; Pittsburgh, PA; and Pawtucket, RI). Participants were assessed at baseline and at 12, 24, 36, 48, 60, 72, 84, and 96 months. The Multicenter Osteoarthritis (MOST) study is also a prospective cohort study of persons who either had knee osteoarthritis or were considered to be at high risk for the development of knee osteoarthritis. The MOST study enrolled 3,026 persons ranging from 50 to 79 years of age from April 2003 to April 2005 at 2 sites (Birmingham, AL; and Iowa City, IA). Participants were assessed at baseline and 15, 30, 60, 72, and 84 months. In both cohorts, no treatment was provided to participants as part of the study, and total knee replacement data were collected at all assessments and were confirmed on the basis of radiographs and/or medical records.

### Total Knee Replacement Timeliness Classification

We applied validated total knee replacement appropriateness criteria at baseline and follow-up time points to both cohorts<sup>17,18</sup>. A procedure is considered appropriate if its expected benefit exceeds its expected negative consequences by a sufficiently wide margin<sup>19</sup>. Developed originally in 2003 by Escobar et al., and updated by Riddle et al., the criteria include 16 mutually exclusive combinations of 5 elements: age, symptomatology, radiographic knee osteoarthritis severity, number of knee compartments involved, and knee stability (Table I)<sup>17,18</sup>. On the basis of these criteria, patients were classified into 3 categories: appropriate, inconclusive, and inappropriate. A description of the criteria is provided in Appendix A; further details have been provided elsewhere<sup>17</sup>. All assessments needed to apply the criteria were available from the MOST dataset. A board-certified musculoskeletal radiologist (C.F.) rated the extent of patellofemoral osteoarthritis for participants in the OAI study, on the basis of magnetic resonance imaging (MRI) findings, with use of a modified Kellgren-Lawrence system as described by Riddle et al.<sup>17</sup>, to determine the number of compartments involved; this information was not readily available in the OAI study. We combined individuals who were classified as “appropriate” and “inconclusive” into 1 group for subsequent analyses as patients with these classifications have been shown to have the same 2-year postoperative trajectories of pain and functional improvement following total knee replacement<sup>20</sup>. Given the preference-based nature of total knee replacement, we defined this group as “potentially appropriate” for total knee replacement and the inappropriate group as “likely inappropriate.”

To qualify for the current study, participants underwent total knee replacement during the follow-up period and/or were classified as “potentially appropriate” either at baseline or at a follow-up visit. Participants were then assigned to 1 of 3 utilization subgroups: (1) “timely” (that is, undergoing total knee replacement within 2 years after being classified as potentially appropriate), (2) “potentially appropriate but not replaced” (that is,

**TABLE II Frequency of Demographic Variables by Total Knee Replacement Utilization Category**

	Timely (Surgery within 2 Years After Becoming Potentially Appropriate) (N = 290, 8%)	Potentially Appropriate but Not Replaced (Surgery Beyond 2 Years After Becoming Potentially Appropriate or No Surgery During Follow-up Period) (N = 2,833, 83%)	Premature (Surgery Likely Inappropriate) (N = 294, 9%)	P Value from Chi-Square Test
Sex*				0.4695
Male	95 (32.76%)	1,002 (35.37%)	96 (32.65%)	
Female	195 (67.24%)	1,831 (64.63%)	198 (67.35%)	
Age*				<0.0001
≤55 yr	35 (12.07%)	434 (15.32%)	68 (23.13%)	
56-63 yr	65 (22.41%)	781 (27.57%)	57 (19.39%)	
64-69 yr	93 (32.07%)	679 (23.97%)	60 (20.41%)	
70-75 yr	61 (21.03%)	602 (21.25%)	62 (21.09%)	
≥76 yr	36 (12.41%)	337 (11.90%)	47 (15.99%)	
Race*				<0.0001
Non-Hispanic white	258 (88.97%)	2,122 (74.90%)	245 (83.33%)	
Black	28 (9.66%)	637 (22.48%)	36 (12.24%)	
Other/missing	4 (1.38%)	74 (2.61%)	13 (4.42%)	
BMI*				<0.0001
<25 kg/m <sup>2</sup>	34 (11.72%)	323 (11.40%)	94 (31.97%)	
25-29 kg/m <sup>2</sup>	89 (30.69%)	902 (31.84%)	78 (26.53%)	
30-34 kg/m <sup>2</sup>	88 (30.34%)	887 (31.31%)	66 (22.45%)	
≥35 kg/m <sup>2</sup>	79 (27.24%)	721 (25.45%)	56 (19.05%)	
Education*				0.0019
High school or less/missing	84 (28.97%)	892 (31.49%)	74 (25.17%)	
Some college	90 (31.03%)	788 (27.82%)	72 (24.49%)	
College degree	46 (15.86%)	492 (17.37%)	46 (15.65%)	
More than college	70 (24.14%)	661 (23.33%)	102 (34.69%)	
Charlson class*				0.0187
0	180 (62.07%)	1,796 (63.40%)	210 (71.43%)	
≥1	110 (37.93%)	1,037 (36.60%)	84 (28.57%)	
Depression (CES-D score, ≥16)*				<0.0001
No	235 (81.03%)	2,298 (81.12%)	270 (91.84%)	
Yes	55 (18.97%)	535 (18.88%)	24 (8.16%)	
SF-12 Physical Function Scale†	37.39 ± 9.37†	39.11 ± 9.26	41.44 ± 9.59†	<0.0001
Living alone*				0.0309
No	251 (86.55%)	2,312 (81.61%)	230 (78.23%)	
Yes	39 (13.45%)	521 (18.39%)	64 (21.77%)	

\*The values are given as the number of patients, with the percentage in parentheses. †The values are given as the mean and the standard deviation. ‡Significant difference between Timely and Premature groups (t test;  $p < 0.05$ ).

being classified as potentially appropriate but not having total knee replacement within the 2-year follow-up period), and (3) “premature” (that is, being classified as “likely inappropriate” at the last follow-up visit before total knee replacement) (Fig. 1). In the absence of a widely accepted definition for timely total knee replacement, the 2-year window following the “potentially appropriate” classification was chosen to allow patients enough time to consider total knee replacement without waiting so long as

to cause further pain and more severe knee osteoarthritis and result in worse postoperative outcomes<sup>21-24</sup>.

#### Analytical Plan

To describe the proportions of patients who were assigned to the 3 utilization subgroups and to derive confidence intervals (CIs), we used as denominators (1) the number of individuals classified as potentially appropriate (for those in the “timely” and

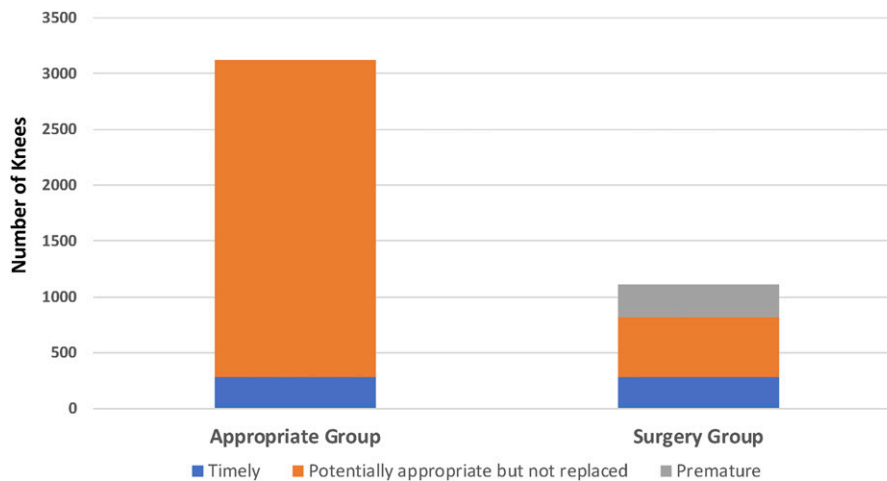


Fig. 2

Histogram showing the proportion of knees characterized as “potentially appropriate but not replaced” among participants for whom total knee replacement was appropriate (left column) and the proportions of knees characterized as “premature” and “potentially appropriate but not replaced” among patients who underwent total knee replacement (right column).

“potentially appropriate but not replaced” subgroups) and (2) the number of total knee replacements (for those in the “premature” subgroup). We then estimated a multinomial logistic mixed-effects regression (MLMR) model with the 3-category outcome (timely, premature, and potentially appropriate but not replaced) and with a site-specific random intercept for the 6 study sites to account for within-site correlation. The variables that were included as fixed effects were age, sex, race, educational status, body mass index (BMI) category, Center for Epidemiological Studies Depression (CES-D) scale score of  $>16$  (indicating clinical depression)<sup>25</sup>, Short Form-12 (SF-12) Physical Component Summary score (PCS)<sup>26</sup>, Charlson Comorbidity Index score (derived from a participant questionnaire)<sup>27</sup>, and whether or not the participant was living alone.

We conducted a number of sensitivity analyses to address the potential for misclassification. First, to address the potential for misclassification as being “potentially appropriate but not replaced,” we applied a more conservative 4-year cutoff (instead of a 2-year cutoff). We also limited our analysis to participants whose classification was consistently “potentially appropriate” over a 2-year period (planned analysis) and 4-year period (sensitivity analysis) from the time they were first classified as such. Finally, we limited our analysis to participants who were classified as “potentially appropriate” by excluding those who were classified as “inconclusive.” Second, to address the potential for misclassification as “premature,” we limited our analyses to participants with no more than 1 year between their appropriateness assessment and total knee replacement to decrease the possibility of deterioration in the knee between the assessment for appropriateness and surgery<sup>28,29</sup> (to minimize the potential overestimation of the likelihood of being classified as “premature”). Third, to address the potential effect of statistical dependence of 2 knees from the same individual, we reran our models by randomly selecting and including only 1 of the knees. The study was approved by our institutional review board.

## Results

Appropriateness classification was applied to both knees of all 8,002 participants (16,004 knees) from the OAI and MOST studies, of whom 852 underwent 1,114 total knee replacements over the study period. Of the examined knees, 12,587 were not studied further because they were classified as “inappropriate” and did not have total knee replacement during the study period. The remaining 3,417 knees (from 2,313 individuals) were classified into 1 of the 3 utilization categories as follows: 3,123 knees (including 1,920 inconclusive knees) were considered to be “potentially appropriate” for total knee replacement (of which 290 knees were classified as “timely” and 2,833 knees were classified as “potentially appropriate but not replaced”), and 294 knees were considered to be “likely inappropriate” yet underwent total knee replacement and thus were classified as “premature”. In the “potentially appropriate but not replaced” group, 1,204 (42.5%) of the 2,833 knees had severe symptoms (that is, combined Western Ontario and McMaster Universities [WOMAC] Pain and Physical Function scores of  $\geq 34$ <sup>17</sup>). Characteristics of the 3 utilization groups are presented in Table II.

The 294 total knee replacements that were classified as “premature” represented 26.4% (95% CI, 23.8% to 29.1%) of the 1,114 total knee replacements performed during the study period (Fig. 2). The 2,833 knees that were classified as “potentially appropriate but not replaced” represented 90.7% (95% CI, 89.6% to 91.7%) of the “potentially appropriate” category ( $n = 3,123$ ) (Fig. 2).

When all 3 utilization groups were included in the MLMR analysis, with “timely” as the reference category, the odds of being “potentially appropriate but not replaced” were 2.8 times greater in blacks compared with whites (Table III). The odds of being “premature” increased with living alone and decreased with being overweight or obese, having depressive symptoms, and older age.

Sensitivity analyses showed that using a 4-year limit for timely use reduced the proportion of those “potentially



**TABLE III Factors Associated with Being Characterized as “Premature” or “Potentially Appropriate but Not Replaced” \*†**

	Potentially Appropriate but Not Replaced Versus Timely	Premature Versus Timely
Sex (reference, female)		
Male	1.17 (0.89, 1.53)	0.97 (0.68, 1.40)
Age (reference, ≤55 years)		
56-63 yr	1.05 (0.68, 1.62)	<b>0.56 (0.32, 0.98)</b>
64-69 yr	<b>0.64 (0.42, 0.98)</b>	<b>0.35 (0.21, 0.61)</b>
70-75 yr	0.87 (0.56, 1.36)	<b>0.53 (0.30, 0.93)</b>
≥76 yr	0.87 (0.53, 1.43)	0.79 (0.43, 1.48)
Race (reference, non-Hispanic white)		
Black	<b>2.81 (1.86, 4.25)</b>	1.68 (0.97, 2.90)
Other/missing	2.32 (0.84, 6.44)	<b>3.56 (1.12, 11.36)</b>
BMI (reference, <25 kg/m <sup>2</sup> )		
25-29 kg/m <sup>2</sup>	1.04 (0.68, 1.58)	<b>0.33 (0.20, 0.55)</b>
30-34 kg/m <sup>2</sup>	1.02 (0.67, 1.56)	<b>0.31 (0.18, 0.52)</b>
≥35 kg/m <sup>2</sup>	0.87 (0.56, 1.36)	<b>0.33 (0.19, 0.56)</b>
Education (reference, high school or less/missing)		
Some college	0.82 (0.59, 1.12)	0.93 (0.59, 1.45)
College degree	1.02 (0.70, 1.50)	1.03 (0.61, 1.74)
More than college	0.91 (0.65, 1.27)	1.50 (0.96, 2.36)
Charlson score (reference, 0)		
≥1	0.98 (0.76, 1.27)	0.79 (0.55, 1.13)
Depressed	0.88 (0.64, 1.21)	<b>0.37 (0.22, 0.64)</b>
SF-12 Physical Function Scale	<b>1.02 (1.01, 1.03)</b>	<b>1.04 (1.02, 1.06)</b>
Living alone (reference, no)		
Yes	1.42 (0.99, 2.05)	<b>2.04 (1.29, 3.22)</b>

\*The values are given as the odds ratio, with the 95% CI in parentheses. †Bold values indicate that the odds ratio is significantly different from 1.

appropriate but not replaced” to 76.2%. Limiting the analysis to participants whose classification was consistently “potentially appropriate” over the initial 2-year period (n = 2,966) resulted in the proportions of “timely” and “potentially appropriate but not replaced” being 9.5% and 80.6%, respectively. With a 4-year period of consistent classification (n = 2,476), the proportions of “timely” and “potentially appropriate but not replaced” were 11.4% and 76.8%, respectively. Excluding individuals with inconclusive classification reduced the percentage of those “potentially appropriate but not replaced” to 72.0%. When the analyses were limited to participants with no more than a 1-year lapse between their last appropriateness assessment and a total knee replacement, the proportion of “premature” was 7.2%. Effect estimates from the sensitivity analysis regressions were relatively unchanged from the main analyses (see Appendix B).

## Discussion

To our knowledge, the present study is the first to quantify the extent of appropriateness for and timeliness of total knee replacement. Most (90.7%) of the knees that were classified as potentially appropriate for total knee replacement were unreplaced within our time frame for timely use, whereas approximately one-quarter (26.4%) of the total knee replacements that were performed appeared to have been done pre-

maturely. These findings were robust in multiple sensitivity analyses. Being of black race was associated with an increased chance of being in the “potentially appropriate but not replaced” subgroup, and younger age, living alone, normal weight, and no evidence of clinical depression were associated with premature use. With nearly 1,000,000 total knee replacements being performed in the U.S. each year<sup>2,3</sup>, the present study provides important impetus for future efforts directed at measuring and improving timely utilization of total knee replacement.

Our study suggests that there is a potentially large number of patients with knee osteoarthritis in the U.S. who could benefit from total knee replacement but do not undergo the procedure. The majority of patients who were potentially appropriate for total knee replacement did not undergo the procedure within 2 to 4 years after it had become appropriate. Among those, nearly half had severe symptoms (that is, combined WOMAC Pain and Physical Function scores of ≥34) and therefore potentially could have experienced large improvement following surgery<sup>1</sup>. While it is reasonable to assume that a subset of such patients have sensible reasons for not undergoing total knee replacement (for example, medical contraindications or the need to delay surgery because of the responsibility of caring for others (e.g., a spouse or parent), the rest actually may benefit from surgery. Black race, which was statistically associated with an increased chance of being in the

“potentially appropriate but not replaced” subgroup, provides 1 such example. In the U.S., blacks are less likely to utilize total knee replacement than whites<sup>7,12,30</sup>, likely because of individual and environmental factors. Blacks are more likely to have lower expectations of surgery<sup>31</sup>. In addition, physician networks in communities with higher proportions of black residents are more likely to be smaller and more insular, which may result in lower utilization of and poorer access to total knee replacement<sup>32</sup>. Understanding the reasons and barriers resulting in delay in those subpopulations is important as the superiority of total knee replacement over non-operative management has been demonstrated<sup>1</sup>.

With 26% of the 1,114 total knee replacements in our study classified as premature, our study also suggests that some patients currently undergoing total knee replacement are doing so prematurely and, based on previous work<sup>20</sup>, likely will derive little benefit from the procedure in terms of pain relief and functional improvement. Given that total knee replacement is a major operation, this observation raises the question of whether these patients are undergoing surgery for some other potential benefits such as fulfillment of social roles, which are neither captured in the appropriateness criteria nor reflected in the outcomes against which they were validated. Our results provide support for these conjectures. Younger age and living alone were associated with premature total knee replacement. The most rapid increase in total knee replacement rates is projected among patients <65 years of age<sup>33,34</sup>, who are still in the workforce and may be undergoing surgery for economic necessities such as the anticipation of work loss or return to work. Similarly, the anticipation of disability associated with knee osteoarthritis may be driving patients who live alone to seek surgery when the disease is in its early stages in order to avoid the functional limitations associated with progression of the disease. Whereas some of those patients may indeed have legitimate reasons for undergoing total knee replacement early, further research is needed to revisit what is appropriate and to disentangle early yet appropriate total knee replacement use from potential overuse.


Our findings illustrate the potential ramifications of applying appropriateness criteria to the utilization of an elective procedure such as total knee replacement. Although appropriateness criteria are conventionally used to identify cases of overuse for cost containment, the present study clearly shows that applying such criteria to populations of patients may result in an overall increase in projected health-care costs as the potential for increased utilization of total knee replacement among potentially appropriate patients with knee osteoarthritis overshadows the savings from restraining premature use. Yet, as alluded to earlier, these findings should be interpreted with caution as the available criteria provide only a rough yardstick that can inform the current patterns of use. Despite the fact that the Escobar appropriateness classifications (that is, appropriate or inappropriate) predict pain and function outcome trajectories<sup>20</sup>, some of the included variables used to determine appropriateness involved arbitrary cutoffs (for example, an age of <55 years as compared with 55 to 65 years and Kellgren-Lawrence grade of 3 compared with 4) that were not evidence-

based, may not be clinically relevant, and may not reflect contemporary practice (for example, judgments of knee stability on the basis of lateral knee joint gapping). The Escobar system was published in 2003, nearly 2 decades ago and likely does not capture more contemporary patient-relevant preferences, particularly those of younger and more-active patients. More refined and contemporaneous appropriateness criteria are needed to more accurately assess total knee replacement utilization in the U.S. The next generation of appropriateness criteria should consider a more encompassing definition of benefit, potentially including data from registries (for example, data on expected improvement) as well as considering social determinants of health (for example, living alone) to evaluate appropriateness. It is our hope that the findings provided in the present report will help to catalyze the development of contemporary criteria.

Our study has important limitations. Although the data were obtained from 6 diverse sites (2 for the MOST study and 4 for the OAI study) that included both university-based institutions and community-based sites, and although we included a site-specific random effect in our analyses to increase generalizability, these sites may not be representative of all locations that treat patients who have knee osteoarthritis. The MOST and OAI studies represent observational cohorts of community subjects with knee osteoarthritis, the majority of whom were not seeing a physician for knee osteoarthritis, and there was no study protocol for referral for evaluation and treatment unless abnormalities such as tumors or osteonecrosis were present on images. Income and insurance may be associated with utilization, yet those data were not consistently collected in the 2 cohorts. Finally, participants with incomplete follow-up were not excluded; however, we addressed this potential limitation in a sensitivity analysis and restricted our sample to the subset with no more than 1 year between symptom assessment and total knee replacement, with similar results.

Limitations notwithstanding, in the absence of national longitudinal data on patients with knee osteoarthritis to whom appropriateness criteria can be applied, these data provide the best available estimates of the timeliness of total knee replacement utilization in the U.S. Further understanding of these observations is needed, especially in the cohort of potentially appropriate patients whose knees remained unreplaced after 2 years.

## Appendix

 Supporting material provided by the authors is posted with the online version of this article as a data supplement at <http://links.lww.com/JBJS/F668>. ■

H.M.K. Ghomrawi, PhD, MPH<sup>1</sup>  
A.I. Mushlin, MD, ScM<sup>2</sup>  
R. Kang, MA<sup>1</sup>  
S. Banerjee, PhD<sup>2</sup>  
J.A. Singh, MD<sup>3</sup>  
L. Sharma, MD<sup>1</sup>  
C. Flink, MD<sup>4</sup>  
M. Nevitt, PhD<sup>5</sup>



T. Neogi, MD, PhD<sup>6</sup>  
D.L. Riddle, PT, PhD<sup>7</sup>

<sup>1</sup>Departments of Surgery (H.M.K.G.), Pediatrics (H.M.K.G.), and Medicine (L.S.), Center for Health Services and Outcomes Research (H.M.K.G. and R.K.), Feinberg School of Medicine of Northwestern University, Chicago, Illinois

<sup>2</sup>Department of Healthcare Policy and Research, Weill Cornell Medical College, New York, NY

<sup>3</sup>Department of Medicine, University of Alabama at Birmingham and Birmingham VA Medical Center, Birmingham, Alabama

<sup>4</sup>Department of Radiology, University of Cincinnati Medical Center, Cincinnati, Ohio

<sup>5</sup>Medical School, University of California, San Francisco, San Francisco, California

<sup>6</sup>Clinical Epidemiology Unit, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts

<sup>7</sup>Departments of Physical Therapy, Orthopedics, and Rheumatology, Virginia Commonwealth University, Richmond, Virginia

Email address for H.M.K. Ghomrawi:  
hassan.ghomrawi1@northwestern.edu

ORCID iD for H.M.K. Ghomrawi: [0000-0002-9686-1010](https://orcid.org/0000-0002-9686-1010)

ORCID iD for A.I. Mushlin: [0000-0002-9298-9084](https://orcid.org/0000-0002-9298-9084)

ORCID iD for R. Kang: [0000-0001-7762-8152](https://orcid.org/0000-0001-7762-8152)

ORCID iD for S. Banerjee: [0000-0001-6936-0846](https://orcid.org/0000-0001-6936-0846)

ORCID iD for J.A. Singh: [0000-0003-3485-0006](https://orcid.org/0000-0003-3485-0006)

ORCID iD for L. Sharma: [0000-0002-5571-1622](https://orcid.org/0000-0002-5571-1622)

ORCID iD for C. Flink: [0000-0003-0680-6868](https://orcid.org/0000-0003-0680-6868)

ORCID iD for M. Nevitt: [0000-0002-6988-7530](https://orcid.org/0000-0002-6988-7530)

ORCID iD for T. Neogi: [0000-0002-9515-1711](https://orcid.org/0000-0002-9515-1711)

ORCID iD for D.L. Riddle: [0000-0002-3611-2739](https://orcid.org/0000-0002-3611-2739)

## References

- Losina E, Walensky RP, Kessler CL, Emrani PS, Reichmann WM, Wright EA, Holt HL, Solomon DH, Yelin E, Paltiel AD, Katz JN. Cost-effectiveness of total knee arthroplasty in the United States: patient risk and hospital volume. *Arch Intern Med*. 2009 Jun 22;169(12):1113-21; discussion 1121-2.
- Maradit Kremers H, Larson DR, Crowson CS, Kremers WK, Washington RE, Steiner CA, Jiranek WA, Berry DJ. Prevalence of total hip and knee replacement in the United States. *J Bone Joint Surg Am*. 2015 Sep 2;97(17):1386-97.
- Kurtz SM, Ong KL, Lau E, Bozic KJ. Impact of the economic downturn on total joint replacement demand in the United States: updated projections to 2021. *J Bone Joint Surg Am*. 2014 Apr 16;96(8):624-30.
- Ghomrawi HM, Schackman BR, Mushlin AI. Appropriateness criteria and elective procedures—total joint arthroplasty. *N Engl J Med*. 2012 Dec 27;367(26):2467-9.
- Vastag B. Knee replacement underused, says panel: useful option when non-surgical therapies fail. *JAMA*. 2004 Jan 28;291(4):413-4.
- Wilson NA, Schneller ES, Montgomery K, Bozic KJ. Hip and knee implants: current trends and policy considerations. *Health Aff (Millwood)*. 2008 Nov-Dec;27(6):1587-98.
- Skinner J, Weinstein JN, Sporer SM, Wennberg JE. Racial, ethnic, and geographic disparities in rates of knee arthroplasty among Medicare patients. *N Engl J Med*. 2003 Oct 2;349(14):1350-9.
- James AE 3rd, Gellad WF, Primack BA. Implications of new insurance coverage for access to care, cost-sharing, and reimbursement. *JAMA*. 2014 Jan 15;311(3):241-2.
- Hanchate AD, Kapoor A, Katz JN, McCormick D, Lasser KE, Feng C, Manze MG, Kressin NR. Massachusetts health reform and disparities in joint replacement use: difference in differences study. *BMJ*. 2015 Feb 20;350:h440.
- Baicker K, Taubman SL, Allen HL, Bernstein M, Gruber JH, Newhouse JP, Schneider EC, Wright BJ, Zaslavsky AM, Finkelstein AN; Oregon Health Study Group. The Oregon experiment—effects of Medicaid on clinical outcomes. *N Engl J Med*. 2013 May 2;368(18):1713-22.
- Albert MA, Ayanian JZ, Silbaugh TS, Lovett A, Resnic F, Jacobs A, Normand SL. Early results of Massachusetts healthcare reform on racial, ethnic, and socioeconomic disparities in cardiovascular care. *Circulation*. 2014 Jun 17;129(24):2528-38. Epub 2014 Apr 11.
- Singh JA, Lu X, Rosenthal GE, Ibrahim S, Cram P. Racial disparities in knee and hip total joint arthroplasty: an 18-year analysis of national Medicare data. *Ann Rheum Dis*. 2014 Dec;73(12):2107-15. Epub 2013 Sep 18.
- Ravi B, Croxford R, Austin PC, Lipscombe L, Bierman AS, Harvey PJ, Hawker GA. The relation between total joint arthroplasty and risk for serious cardiovascular events in patients with moderate-severe osteoarthritis: propensity score matched landmark analysis. *BMJ*. 2013 Oct 30;347:f6187.
- Turkiewicz A, Kiadaliri AA, Englund M. Cause-specific mortality in osteoarthritis of peripheral joints. *Osteoarthritis Cartilage*. 2019 Jun;27(6):848-54. Epub 2019 Feb 21.
- Kane RL, Wilt T, Suarez-Almazor ME, Fu SS. Disparities in total knee replacements: a review. *Arthritis Rheum*. 2007 May 15;57(4):562-7.
- Ghomrawi HM, Alexiades M, Pavlov H, Nam D, Endo Y, Mandl LA, Mushlin AI. Evaluation of two appropriateness criteria for total knee replacement. *Arthritis Care Res (Hoboken)*. 2014 Nov;66(11):1749-53.
- Riddle DL, Jiranek WA, Hayes CW. Use of a validated algorithm to judge the appropriateness of total knee arthroplasty in the United States: a multi-center longitudinal cohort study. *Arthritis Rheumatol*. 2014 Aug;66(8):2134-43.
- Escobar A, Quintana JM, Aróstegui I, Azkárte J, Güenaga JI, Arenaza JC, Garai I. Development of explicit criteria for total knee replacement. *Int J Technol Assess Health Care*. 2003 Winter;19(1):57-70.
- Lawson EH, Gibbons MM, Ko CY, Shekelle PG. The appropriateness method has acceptable reliability and validity for assessing overuse and underuse of surgical procedures. *J Clin Epidemiol*. 2012 Nov;65(11):1133-43.
- 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 Mar;67(3):349-57.
- Ackerman IN, Bennell KL, Osborne RH. Decline in health-related quality of life reported by more than half of those waiting for joint replacement surgery: a prospective cohort study. *BMC Musculoskelet Disord*. 2011 May 23;12:1108.
- Desmeules F, Dionne CE, Belzile E, Bourbonnais R, Frémont P. The burden of wait for knee replacement surgery: effects on pain, function and health-related quality of life at the time of surgery. *Rheumatology (Oxford)*. 2010 May;49(5):945-54. Epub 2010 Feb 8.
- Lingard EA, Katz JN, Wright EA, Sledge CB, Kinemax Outcomes Group. Predicting the outcome of total knee arthroplasty. *J Bone Joint Surg Am*. 2004 Oct;86(10):2179-86.
- Vina ER, Hannon MJ, Kwok CK. Improvement following total knee replacement surgery: exploring preoperative symptoms and change in preoperative symptoms. *Semin Arthritis Rheum*. 2016 Apr;45(5):547-55. Epub 2015 Oct 22.
- Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas*. 1977 Jun 1;1(3):385-401.
- Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, Bullinger M, Kaasa S, Leplege A, Prieto L, Sullivan M. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQLA Project. *International Quality of Life Assessment*. *J Clin Epidemiol*. 1998 Nov;51(11):1171-8.
- Chaudhry S, Jin L, Meltzer D. Use of a self-report-generated Charlson Comorbidity Index for predicting mortality. *Med Care*. 2005 Jun;43(6):607-15.
- Driban JB, Stout AC, Lo GH, Eaton CB, Price LL, Lu B, Barbe MF, McAlindon TE. Best performing definition of accelerated knee osteoarthritis: data from the Osteoarthritis Initiative. *Ther Adv Musculoskelet Dis*. 2016 Oct;8(5):165-71. Epub 2016 Jul 6.
- Bartlett SJ, Ling SM, Mayo NE, Scott SC, Bingham CO 3rd. Identifying common trajectories of joint space narrowing over two years in knee

osteoarthritis. *Arthritis Care Res (Hoboken)*. 2011 Dec;63(12):1722-8.

**30.** Zhang W, Lyman S, Boutin-Foster C, Parks ML, Pan TJ, Lan A, Ma Y. Racial and ethnic disparities in utilization rate, hospital volume, and perioperative outcomes after total knee arthroplasty. *J Bone Joint Surg Am*. 2016 Aug 3; 98(15):1243-52.

**31.** Groeneveld PW, Kwok CK, Mor MK, Appelt CJ, Geng M, Gutierrez JC, Wessel DS, Ibrahim SA. Racial differences in expectations of joint replacement surgery outcomes. *Arthritis Rheum*. 2008 May 15;59(5):730-7.

**32.** Ghomrawi HMK, Funk RJ, Parks ML, Owen-Smith J, Hollingsworth JM. Physician referral patterns and racial disparities in total hip replacement: a network analysis approach. *PLoS One*. 2018 Feb 20;13(2):e0193014.

**33.** Losina E, Katz JN. Total knee arthroplasty on the rise in younger patients: are we sure that past performance will guarantee future success? *Arthritis Rheum*. 2012 Feb;64(2):339-41.

**34.** Kurtz SM, Lau E, Ong K, Zhao K, Kelly M, Bozic KJ. Future young patient demand for primary and revision joint replacement: national projections from 2010 to 2030. *Clin Orthop Relat Res*. 2009 Oct;467(10):2606-12. Epub 2009 Apr 10.