ACR Open Rheumatology
Vol. 3, No. 9, September 2021, pp 583–592
DOI 10.1002/acr2.11241

American College of Rheumatology
Empowering Rheumatology Professionals

Impact of Preoperative and Incident Musculoskeletal Problematic Areas on Postoperative Outcomes After Total Knee Replacement

MaryAnn Zhang,[1] Faith Selzer,[2] Elena Losina,[2] Jamie E. Collins,[2] and Jeffrey N. Katz[2]

Objective. To examine impact of pre-existing and incident problematic musculoskeletal (MSK) areas after total knee replacement (TKR) on postoperative 60-month Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain/function scores.

Methods. Using data from a randomized controlled trial of subjects undergoing TKR for osteoarthritis, we assessed problematic MSK areas in six body regions before TKR and 12, 24, 36, and 48 months after TKR. We defined the following two variables: *1*) density count (number of problematic MSK areas occurring after TKR; range 0-24) and *2*) cumulative density count (problematic MSK areas both before and after TKR, categorized into four levels: no preoperative areas and density count of 0-1 [reference group]; no preoperative areas and density count of 2 or more; one or more preoperative areas and density count of 0-1; and one or more preoperative areas and density count of 2 or greater). We evaluated the associations between categorized 60-month WOMAC and cumulative density count by ordinal logistic regression.

Results. Among 230 subjects, 24% reported one or more preoperative problematic MSK area. After TKR, 75% reported a density count of 0 to 1; 25% reported a density count of 2 or more. Compared with the reference group, each cumulative density count category was associated with an increased odds of having a higher category of 60-month WOMAC pain score, as follows: 2.97 (95% confidence interval [CI], 1.48-5.98) for no preoperative problematic areas and density count of 2 or greater, 3.31 (95% CI, 1.64-6.66) for one or more preoperative problematic areas and density count of 0 to 1, and 2.85 (95% CI, 0.97-8.39) for one or more preoperative problematic areas and density count of 2 or greater. Similar associations were observed with 60-month WOMAC function score.

Conclusion. In TKR recipients, the presence of problematic musculoskeletal areas beyond the index knee—preoperatively and/or postoperatively—was associated with worse 60-month WOMAC pain/function score.

INTRODUCTION

Knee osteoarthritis affects roughly 14 million people in the United States (1). For those who experience significant knee pain and impaired mobility and who fail conservative management such as analgesics and physical therapy, total knee replacement (TKR) is an important treatment option (2–4). TKR is generally successful at relieving knee pain, although studies report that 20% to 40% of TKR recipients experience some degree of persistent knee pain even 4 years following surgery (5,6). Given the risks of persistent pain and the lengthy rehabilitation process, evaluating an individual's potential for persistent knee pain following surgery may be useful for optimizing TKR outcomes.

One factor to consider when assessing TKR outcomes is concomitant musculoskeletal (MSK) joint complaints beyond the index knee. Several studies have examined the relationship between preoperative MSK complaints and TKR outcomes. Preoperative total Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) scores in the contralateral knee were found to be modestly correlated with worse WOMAC total scores

Supported by National Institute of Arthritis and Musculoskeletal and Skin Diseases grants T32-AR055885 (to MZ), K24-AR 057827 (to EL), and P30-AR072577 (to EL, JNK, FS, and JEC) and a Rheumatology Research Foundation Investigator Award (to JEC).

[1]MaryAnn Zhang, MD, MPH: New York Presbyterian Hospital, Columbia University College of Physicians and Surgeons; [2]Faith Selzer, PhD, Elena Losina, PhD, Jamie E. Collins, PhD, Jeffrey N. Katz, MD, MSc: Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts.

Dr. Losina has received funding from research grants to the Brigham and Women's Hospital from Genentech, Pfizer, and Samumed; has served as a consultant for Pfizer (less than $10 000); and serves as The Journal of Bone and Joint Surgery (JBJS) Deputy Editor for Methodology and Statistics. Dr. Katz has received research grants to the Brigham and Women's Hospital from Samumed and Flexion Therapeutics. He serves as a JBJS Deputy Editor for Methodology and Statistics. Dr. Collins serves as a consultant to Boston Imaging Core Labs (less than $10 000) and serves as Osteoarthritis and Cartilage Associate Editor for *Statistics*. No other disclosures relevant to this article were reported.

Address correspondence to Jeffrey N. Katz MD, MSc, Orthopaedic and Arthritis Center for Outcomes Research, Department of Orthopedic Surgery, 60 Fenwood Road, Suite 5016, Boston, MA 02115. Email: jnkatz@bwh.harvard.edu.

Submitted for publication July 15, 2020; accepted in revised form February 2, 2021.

in the index knee ($r = 0.34$; $P < 0.001$) roughly 3 years following TKR (7). Similarly, preoperative complaints in the neck, ankles, feet, and toes were associated with worse post-TKR WOMAC pain scores (8). Pre-existing lower back pain was also found to lead to worse postoperative pain function and satisfaction scores (9). Together, these studies illustrate that preoperative MSK complaints beyond the index knee may be associated with persistent pain in the operative knee years after TKR.

Far less is known about the relationship between postoperative MSK joint complaints and TKR outcomes. In a previously published study using the same study cohort, Zhang et al reported that nearly half of subjects developed incident problematic MSK areas in the 4 years following TKR, most commonly in the nonindex knee and back (10). Recipients who were female and obese and who had multiple medical comorbidities, worse baseline index knee pain scores, and feelings of anxiety/depression were at higher risk of developing new MSK symptoms after TKR (10). The impact of these incident problematic MSK areas on long-term TKR outcomes has not been examined. If incident MSK areas do negatively affect postoperative outcomes, this information would be valuable to clinicians. Such information would help identify optimal candidates for TKR and guide preoperative counseling to minimize postoperative patient dissatisfaction.

In this study, we sought to determine whether problematic MSK areas (preoperative and incident [postoperative]) after TKR for knee osteoarthritis have an impact on 60-month post-TKR WOMAC pain and function scores beyond that explained by baseline characteristics. We hypothesized that even after adjustment for baseline covariates, the presence of problematic MSK areas before TKR, combined with the development of incident areas after TKR, would be associated with worse 60-month WOMAC pain and function scores.

PATIENTS AND METHODS

Study sample. Our study cohort was drawn from the Adding Value in Knee Arthroplasty (AViKA) Postoperative Care Navigation Trial, a randomized controlled trial that examined the efficacy of a motivational interviewing-based postoperative care navigator compared with usual care to improve functional status after TKR for knee osteoarthritis (11,12). The trial was conducted at Brigham and Women's Hospital (BWH) in Boston, Massachusetts, from 2010 to 2013, with 5-year follow-up data collected through the spring of 2019. All subjects spoke English, were at least 40 years old at the date of TKR, and underwent primary TKR for knee osteoarthritis. We excluded subjects with psychological issues precluding surgery, dementia, nursing home residency, implantation of unicompartmental knee arthroplasty, or bilateral TKR (simultaneous, staged, or planned within six months). The AViKA Navigator Trial was registered at ClinicalTrials.gov (NCT01540851) and was approved by the BWH Institutional Review Board. For this analysis, we included participants with a baseline questionnaire completed within 6 weeks before surgery and who completed the 60-month WOMAC follow-up questionnaire.

Outcome measures. The primary outcomes of this analysis were WOMAC pain and WOMAC function scores measured at the 60-month time point. The WOMAC is a validated questionnaire for hip and knee osteoarthritis that assesses lower extremity pain and function with a five-item pain subscale and 17-item functional limitation subscale (13). We transformed WOMAC scores to a 0- to 100-point scale, with 100 indicating the worst score. We then categorized the 60-month WOMAC scores into 0, 1 to 19, and 20 or greater, corresponding with none, mild, and moderate to severe pain/functional limitation, respectively. We explicitly chose 20 as the WOMAC cutoff score as it equated to "mild" symptoms while ensuring enough subjects in each WOMAC score category for an interpretable analysis.

Predictors of interest. *Definition of problematic MSK area.* We defined "problematic MSK areas" as body regions that were impacted during independent activities of daily living. To measure the presence of problematic MSK areas preoperatively, participants were asked, "In the past four weeks, to what extent did problems in the following areas limit your activities?" Participants could rate "none," "a little," or "a lot" for each of the following areas: neck, hands/wrists/arms/shoulders, back, hips, right knee, left knee, and ankles/feet. We then examined the associations between the following three different approaches to specifying the three-level responses to questions about limitations in body areas: combining "none" and "a little" limitation; combining "a little" and "a lot" of limitation; and assigning one point to "a little" and two points to "a lot" of limitation. Combining "none" and "a little" and summing across baseline, 12-, 24-, 36-, and 48-month responses yielded the highest odds ratio in predicting 60-month WOMAC pain/function scores and, therefore, was the specification used in this analysis. Thus, only responses of "a lot" were deemed clinically relevant and categorized as "presence of a problematic MSK area," with responses of "none" or "a little" viewed as the absence of a problematic MSK area. Excluding the index knee, problematic areas were summed to obtain the total number of MSK problematic areas preoperatively and categorized into zero, one, and two or more areas on the basis of the distribution of these data. In the instance that subjects did not endorse whether a particular MSK-related body site was problematic, the body area was coded as nonproblematic.

Density count. To understand the impact of incident problematic MSK areas on postsurgical outcomes, we created the variable "density count." Density count was defined as the total number of new problematic MSK areas from baseline across any body region after TKR over four follow-up time points (12, 24, 36, and 48 months). To ensure a prospective analysis, we did not include problematic MSK areas at 60 months. To qualify as a "new" problematic MSK area, the area had to be noted as either "none" or "a little" at baseline and "a lot" on at least one follow-up time point. Each instance that an area was rated as "none" or "a little" at baseline and was rated "a lot" at follow-up counted toward the

density count. Given the six body areas and four follow-up time points, the maximum possible density count was 24, but this depended on baseline number of problematic areas. Density count was further categorized into zero to one, two to four, and five or more areas; this categorization followed the skewed distribution of the density count, with the last (five or more) category capturing the long right tail of the distribution. If subjects did not endorse whether a particular MSK-related body site was problematic at baseline, we coded it as nonproblematic (ie, imputed as zero). If a subject returned a follow-up questionnaire but did not respond to individual body site questions, the unanswered site at that particular time point was assumed to be nonproblematic. If a subject did not return a particular year's questionnaire, all body sites at that particular time point were assumed to be nonproblematic.

Cumulative density count. To understand the impact of preoperative and incident problematic MSK areas on postsurgical outcomes, we created the variable "cumulative density count." This variable consisted of the following two components: *1*) the total number of problematic MSK areas preoperatively and *2*) the density count. Cumulative density count was initially categorized into the following four possible levels: *1*) no problematic MSK areas preoperatively and a density count of 0 to 1 (reference group), *2*) no problematic MSK areas preoperatively and a density count of 2 or more, *3*) one or more problematic MSK areas preoperatively and a density count of 0 to 1, and *4*) one or more problematic MSK areas preoperatively and a density count of 2 or more. The number of problematic MSK areas preoperatively was dichotomized at zero versus one or more to compare those with and without preoperative problematic areas. Density count was dichotomized at 0 to 1 versus 2 or more on both clinical and distributional grounds, as we did not view a single problematic area over four years as clinically meaningful, and 76% of subjects had a density count of 0 to 1.

Additional covariates. Baseline covariates included age, sex, body mass index (BMI; in kg/m^2), Charlson Comorbidity Index score (CCI), baseline WOMAC pain and function scores, Mental Health Inventory Five-Item scale (MHI-5), and Pain Catastrophizing Scale (PCS). Age, BMI, and baseline WOMAC pain and function scores were analyzed as continuous variables. The CCI was calculated on the basis of medical comorbidities that were extracted through electronic medical record review and was categorized as 0, 1, and 2 or more; this distribution was the closest to creating roughly equal terciles (14). The MHI-5 is a validated five-item screening tool for anxiety and depression, with scores summed and transformed linearly to a 0- to 100-point scale, with 0 indicating the worst mental health status. On the basis of established cutoff criteria, we further dichotomized MHI-5 scores into less than 68 and 68 or more, in which less than 68 identified subjects at risk of clinically meaningful anxiety and depression (15–17). The PCS, a 13-item score, measures subjects' catastrophizing, or amplified negative cognitions, towards pain, including magnification, rumination, and hopelessness (18,19). Although the PCS is a validated measure, a variety of

Table 1. Baseline characteristics of included versus excluded participants

Characteristics	Included Participants (N = 230)	Excluded Participants (N = 79)	*P* Score
Age, mean (SD), y	66 (7.8)	68 (9.1)	0.04
Sex, n (%)			
Male	85 (37)	38 (48)	0.09
Female	145 (63)	41 (52)	
Body mass index, mean (SD), kg/m^2	31 (5.8)	31 (7.0)	0.74
Charlson Comorbidity Index, n (%)			
0	101 (47)	29 (40)	0.28
1	42 (19)	11 (15)	
≥2	74 (34)	32 (44)	
Number of musculoskeletal problematic areas preoperatively[a], n (%)			
0	174 (76)	54 (70)	0.44
1	37 (16)	13 (17)	
≥2	19 (8)	10 (13)	
Baseline WOMAC pain score,[b] mean (SD)	40 (17)	42 (20)	0.35
Baseline WOMAC function score, mean (SD)	40 (16)	43 (19)	0.18
Mental Health Inventory Five-Item Score,[c] n (%)			
<68	40 (17)	26 (34)	0.004
≥68	190 (83)	51 (66)	
Pain Catastrophizing Score,[d] n (%)			
<16	177 (78)	49 (66)	0.05
≥16	50 (22)	25 (34)	

WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.
[a] Presence of preoperative musculoskeletal problematic area was defined as a response of "a lot of limitation" for at least one region on baseline questionnaire.
[b] Scores were transformed to a 0- to 100-point scale (100 indicating the worst score).
[c] Scores were transformed to a 0- to 100-point scale (0 indicating the worst score).
[d] Scores were dichotomized so that 16 or more was considered high pain catastrophizing.

cutoff points have been cited across the literature. The distribution of PCS scores in our cohort was skewed toward 0, with 76% of subjects reporting a PCS score of less than 16. As a result, we dichotomized PCS into less than 16 or 16 or more, in which 16 or more equated to high pain catastrophizing, as this cutoff correlated with the 75th percentile of PCS scores in our cohort. A similar distributional approach was employed by the creators of the PCS for establishing cutoff scores (18). Another study examining preoperative pain catastrophizing and post-TKR outcomes also used a score of 16 as the cutoff (18,19). Although 16 or more may be an appropriate cutoff score in our study, its clinical relevance is unclear, and this cutoff may not be generalizable to other studies.

Given that the AViKA trial was negative with respect to the association between intervention and 6-month WOMAC pain score (primary aim), we did not adjust for the intervention arm.

Statistical analysis. We present descriptive statistics for the overall cohort as means (SDs) for continuous variables and as counts (percentages) for categorical variables. When comparing participants who completed the 60-month follow-up questionnaire with those who did not, we compared continuous variables by Student's t-tests or Wilcoxon rank sum test depending on distribution normality, whereas categorical variables were compared by χ^2 or Fisher's exact test (Table 1). We assessed the association between 60-month WOMAC pain and function scores (categorized as 0, 1-19, and ≥20) with baseline characteristics using χ^2 or Fisher's exact test for categorical variables and ANOVA or Kruskal-Wallis test for continuous variables.

Primary analysis. To assess whether preoperative and incident postoperative problematic MSK areas influenced 60-month WOMAC pain and function categories, we used ordinal logistic regression, with the outcome (ie, 60-month WOMAC pain score) categorized at 0, 1 to 19, and 20 or more. We elected to use this method given the highly skewed nature of the 60-month WOMAC scores. The cut points are clinically meaningful; a score of 0 denotes no pain, and a score of 20 indicates that, on average, the subjects endorsed "mild" pain on the WOMAC pain items. We created the following three models for each outcome: *1*) a baseline model, *2*) the baseline model with the addition of the cumulative density count (divided into four levels), and *3*) the baseline model with cumulative density count (dichotomized). To create the baseline model, we included age (per 5 years) and sex in all models, as we deemed these to be clinically relevant. We then used stepwise selection for all other covariates (BMI, CCI, baseline WOMAC pain or function score per 10 points, MHI-5 <68, and PCS ≥16), with entry and exit criteria of *P* of less than 0.05. For the third model, we dichotomized cumulative density county by keeping the same reference group and collapsing the remaining three levels into a single indicator (≥1 problematic MSK areas preoperatively or a density count of ≥2) because the remaining three levels produced similar odds ratios in model 2. We calculated odds ratios (ORs), 95% confidence intervals (CIs), and C-statistics for all models. All models met the proportional odds assumption.

Sensitivity analysis. To assess whether density count alone influenced 60-month WOMAC pain and function scores, we created the following two ordinal logistic regression models for each WOMAC outcome: *1*) a baseline model and *2*) the baseline model with the addition of the density count. The same methodology as described above was used to determine which covariates to include in the model. These models included "number of problematic MSK areas at baseline" and density count as separate covariates. We calculated ORs, 95% CIs, and C-statistics for all models. All models met the proportional odds assumption.

Analyses were conducted using SAS 9.4 (SAS Institute). Reported *P* values are two-sided, and *P* of less than 0.05 was considered statistically significant.

RESULTS

Patient characteristics. Baseline characteristics of the cohort are displayed in Table 2. Among the 309 enrolled subjects, 230 (74%) completed the baseline and 60-month follow-up questionnaires and were included in the analyses. Of these 230, two subjects completed the 60-month WOMAC function score but failed to complete the 60-month WOMAC pain score. As a

Table 2. Baseline Characteristics of Cohort

Characteristics	Results (N = 230)
Age, mean (SD), y	66 (7.8)
Sex, n (%)	
Male	85 (37)
Female	145 (63)
Body mass index, mean (SD), kg/m^2	31 (5.8)
Charlson Comorbidity Index, n (%)	
0	101 (47)
1	42 (19)
≥2	74 (34)
Number of musculoskeletal problematic areas preoperatively,[a] n (%)	
0	174 (76)
1	37 (16)
≥2	19 (8)
Baseline WOMAC pain score,[b] mean (SD)	40 (17)
Baseline WOMAC function score, mean (SD)	40 (16)
Mental Health Inventory 5-Item Score,[c] n (%)	
<68	40 (17)
≥68	190 (83)
Pain Catastrophizing Score,[d] n (%)	
<16	177 (78)
≥16	50 (22)

WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.
Individuals who did not complete the 60-month follow-up questionnaire were excluded (n = 79).
[a] Presence of preoperative musculoskeletal problematic area was defined as a response of "a lot of limitation" for at least one region on baseline questionnaire.
[b] Scores were transformed to a 0- to 100-point scale (100 indicating the worst score).
[c] Scores were transformed to a 0- to 100-point scale (0 indicating the worst score).
[d] Scores were dichotomized so that 16 or more was considered high pain catastrophizing.

result, all analyses with 60-month WOMAC pain score as the outcome included only 228 subjects. Of the 230 subjects, three (1%) completed only one follow-up questionnaire, five (2%) completed two follow-up questionnaires, 11 (5%) completed three follow-up questionnaires, 27 (12%) completed four follow-up questionnaires, and 184 (80%) completed all five follow-up questionnaires. For any particular MSK body region at any time point, roughly 15% of responses were missing and thus imputed.

Baseline characteristics of included versus excluded participants are displayed in Table 1. At baseline, excluded participants (n = 79) were slightly older (68 years old among the excluded versus 66 years old among the included) and more likely to report feelings of anxiety and depression (MHI-5 scores of <68) (34% of the excluded versus 17% of the included) compared with subjects in the analytic cohort. There were no clinically important or statistically significant differences between included and excluded participants with respect to all other characteristics.

Of the 230 subjects included in this analysis, the mean (SD) age was 66 (7.8) years, 63% were female, the mean (SD) BMI was 31 (5.8) kg/m^2, and 66% had a CCI of 0 or 1. With regard to presence of nonindex knee problematic MSK areas preoperatively, 76% had none, 16% had one area involved, and 8% had two or more areas involved. The mean (SD) baseline WOMAC score was 40 (17) for pain and 40 (16) for function. Seventeen percent of participants reported substantial baseline symptoms of anxiety and depression (MHI-5 score of <68), whereas 22% of participants demonstrated high pain catastrophizing (PCS ≥16).

Sixty-month WOMAC pain and function scores by baseline characteristics, density count, and cumulative density count. Table 3 displays the distribution of 60-month WOMAC pain scores (categorized as 0, 1-19, and ≥20) by baseline characteristics, density count, and cumulative density count. Significant differences in the distribution of 60-month WOMAC pain scores were observed with the following characteristics: number

Table 3. Sixty-month WOMAC pain scores by baseline characteristics, density count, and cumulative density count

	60-Month WOMAC Pain Scores			
	0 (N = 127)	1-19 (N = 61)	≥20 (N = 40)	*P* Value
Age, mean (SD), y	66 (7.8)	66 (7.9)	66 (7.6)	0.91
Sex, n (%)				
Male	56 (66)	19 (22)	10 (12)	0.05
Female	71 (50)	42 (29)	30 (21)	
Body mass index, mean (SD), kg/m^2	30 (5.7)	32 (6.0)	30 (5.6)	0.37
Charlson Comorbidity Index, n (%)				
0	59 (59)	23 (23)	18 (18)	0.6
1	19 (45)	14 (33)	9 (21)	
≥2	43 (59)	18 (25)	12 (16)	
Number of musculoskeletal problematic areas preoperatively, n (%)				
0	106 (61)	43 (25)	24 (14)	0.004
1	17 (46)	12 (32)	8 (22)	
≥2	4 (22)	6 (33)	8 (44)	
Baseline WOMAC pain score,[a] mean (SD)	38 (16)	39 (17)	50 (17)	0.0001
Baseline WOMAC function score, mean (SD)	38 (16)	40 (16)	49 (15)	0.0006
Mental Health Inventory Five-Item Score,[b] n (%)				
<68	12 (31)	13 (33)	14 (36)	0.0005
≥68	115 (61)	48 (25)	26 (14)	
Pain Catastrophizing Score,[c] n (%)				
<16	104 (59)	47 (27)	26 (15)	0.17
≥16	22 (46)	14 (29)	12 (25)	
Density count, n (%)				
0-1	106 (61)	45 (26)	22 (13)	0.0009
2-4	16 (37)	15 (35)	12 (28)	
≥5	5 (42)	1 (8)	6 (50)	
Cumulative density count, n (%)				
0 preoperative areas/0-1 density count	89 (67)	33 (25)	10 (8)	<0.0001
0 preoperative areas/≥2 density count	17 (41)	10 (24)	14 (34)	
≥1 preoperative areas/0-1 density count	17 (41)	12 (29)	12 (29)	
≥1 preoperative areas/ ≥2 density count	4 (29)	6 (43)	4 (29)	

WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.
[a] Scores were transformed to a 0- to 100-point scale (100 indicating the worst score).
[b] Scores were transformed to a 0- to 100-point scale (0 indicating the worst score).
[c] Scores were dichotomized so that 16 or more was considered high pain catastrophizing.

of MSK problematic areas preoperatively, baseline WOMAC pain score, baseline WOMAC function score, MHI-5, density count, and cumulative density count. Table 4 displays the distribution of 60-month WOMAC function scores (categorized as 0, 1-19, and ≥20) by preoperative characteristics, density count, and cumulative density count. Significant differences in the distribution of 60-month WOMAC function scores were observed with the following characteristics: number of MSK problematic areas at baseline, baseline WOMAC pain score, baseline WOMAC function score, MHI-5, and cumulative density count. Density count, however, was not significantly different between 60-month WOMAC function score groups.

Primary analysis: ordinal logistic regression by cumulative density counts. The ordinal logistic regression model for 60-month WOMAC pain score showed significant associations between pain and cumulative density count (model 2, Table 5). Compared with those with no preoperative problematic MSK areas and a density count of 0 to 1, the adjusted OR for cumulative density count was on the order of 3.0 for the other groups: those with no preoperative problematic MSK areas and a density count of 2 or more, those with one or more preoperative problematic MSK area and a density count of 0 to 1, and those with both one or more preoperative problematic MSK area and a density count of 2 or more. The model C-statistic for the initial model (model 1) was 0.64; it improved to 0.70 (model 2), a relative increase of 9%, with the addition of cumulative density count. Because the model showed similar ORs for subjects who had one or more preoperative problematic area, a density count of 2 or more, or both, we combined these variables into a single indicator, as follows: no preoperative problematic areas and density count 0 to 1 versus one or more preoperative problematic area or density count of 2 or more. In this model, as compared with those with no preoperative problematic areas and density count 0 to 1, the adjusted

Table 4. Sixty-month WOMAC function scores by baseline characteristics, density count, and cumulative density count

	60-Month WOMAC Function Scores			
	0 (N = 70)	1-19 (N = 114)	≥20 (N = 46)	*P* Value
Age, mean (SD), y	65 (8.4)	66 (7.6)	67 (7.2)	0.41
Sex, n (%)				
Male	30 (35)	43 (51)	12 (14)	0.18
Female	40 (28)	71 (49)	34 (23)	
Body mass index, mean (SD), kg/m^2	30 (5.7)	31 (5.8)	31 (5.7)	0.47
Charlson Comorbidity Index, n (%)				
0	34 (34)	47 (47)	20 (20)	0.92
1	13 (31)	21 (50)	8 (19)	
≥2	20 (27)	38 (51)	16 (22)	
Number of musculoskeletal problematic areas at baseline, n (%)				
0	60 (34)	89 (51)	25 (14)	0.0001
1	9 (24)	18 (49)	10 (27)	
≥2	1 (5)	7 (37)	11 (58)	
Baseline WOMAC pain score,[a] mean (SD)	39 (16)	38 (16)	48 (19)	0.0014
Baseline WOMAC function score, mean (SD)	38 (16)	39 (16)	48 (16)	0.0016
Mental Health Inventory Five-Item Score,[b] n (%)				
<68	5 (13)	20 (50)	15 (38)	0.0019
≥68	65 (34)	94 (49)	31 (16)	
Pain Catastrophizing Score,[c] n (%)				
<16	58 (33)	90 (51)	29 (16)	0.07
≥16	11 (22)	24 (48)	15 (30)	
Density count, n (%)				
0-1	60 (35)	84 (49)	29 (17)	0.09
2-4	8 (18)	24 (53)	13 (29)	
≥5	2 (17)	6 (50)	4 (33)	
Cumulative density count, n (%)				
0 baseline areas/0-1 density count	51 (39)	68 (52)	13 (10)	0.0002
0 baseline areas/≥2 density count	9 (21)	21 (50)	12 (29)	
≥1 baseline areas/0-1 density count	9 (22)	16 (39)	16 (39)	
≥1 baseline areas/ ≥2 density count	1 (7)	9 (60)	5 (33)	

WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.
[a] Scores were transformed to a 0- to 100-point scale (100 indicating the worst score).
[b] Scores were transformed to a 0- to 100-point scale (0 indicating the worst score).
[c] Scores were dichotomized so that 16 or more was considered high pain catastrophizing.

Table 5. Ordinal logistic regression models for 60-month WOMAC pain category by cumulative density count

Variable	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Age per 5 y	1.04 (0.88-1.23)	0.99 (0.83-1.17)	0.99 (0.83-1.18)
Female sex	2.16 (1.24-3.77)	1.88 (1.06-3.35)	1.86 (1.05-3.29)
MHI-5 <68	3.78 (1.95-7.33)	3.19 (1.61-6.31)	3.14 (1.60-6.18)
Cumulative density count[a]			
0 preoperative areas/0-1 density count	-	Ref	Ref
0 preoperative areas/≥2 density count	-	2.97 (1.48-5.98)	-
≥1 preoperative areas/0-1 density count	-	3.31 (1.64-6.66)	-
≥1 preoperative areas/ ≥2 density count	-	2.85 (0.97-8.38)	-
≥1 preoperative areas or ≥2 density count[b]	-	-	3.10 (1.79-5.35)
C-statistic	0.64	0.70	0.69

CI, confidence interval; MHI-5, Mental Health Inventory (0 indicating the worst score); OR, odds ratio; Ref, reference; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index (categorized as 0, 1-19, and ≥20).
The following covariates were also included in the models but did not reach statistical significance: body mass index, Charlson Comorbidity Index, baseline WOMAC pain score, and Pain Catastrophizing Scale.
[a]Cumulative density count consisted of number of problematic musculoskeletal areas present preoperatively (either 0 or ≥1) and density count (either 0-1 or ≥2).
[b]In model 3 only, cumulative density count consisted of the following two groups: *1*) zero preoperative areas and 0 to 1 density count and *2*) one or more preoperative area or 2 or greater density count.

OR for the presence of one or more preoperative problematic area or density count of 2 or more was 3.10 (95% CI, 1.79-5.35) (model 3, Table 5).

Based on the ordinal logistic regression model for 60-month WOMAC function (model 2, Table 6), the adjusted OR for cumulative density count was similar to those observed with 60-month WOMAC pain score, with ORs for each category of the cumulative density count ranging from 2.17 to 3.54 compared with those with no problematic MSK areas preoperatively and a density count of 0 to 1. The C-statistic for the initial model (model 1) was 0.64; it improved to 0.68 in the final model (model 2). As with the WOMAC pain model, we then combined the MSK pain variables into a single indicator. In this model and relative to those without preoperative MSK pain and who reported 0 to 1 painful areas during follow-up, the adjusted OR for the presence of one or more preoperative problematic area or density count of 2 or more was 2.80 (95% CI, 1.62-4.81) (model 3, Table 6).

Table 6. Ordinal logistic regression models for 60-month WOMAC function category by cumulative density count

Variable	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Age per 5 y	1.17 (0.99-1.38)	1.10 (0.93-1.29)	1.10 (0.94-1.30)
Female sex	1.53 (0.90-2.58)	1.46 (0.86-2.49)	1.45 (0.85-2.45)
Baseline WOMAC function per 10 points	1.20 (1.03-1.41)	-	-
MHI-5 <68	3.08 (1.55-6.12)	3.10 (1.55-6.19)	2.97 (1.50-5.89)
Cumulative density count[a]			
0 preoperative areas/0-1 density count	-	Ref	Ref
0 preoperative areas/≥2 density count	-	2.17 (1.08-4.36)	-
≥1 preoperative areas/0-1 density count	-	3.54 (1.74-7.19)	-
≥1 preoperative areas/ ≥2 density count	-	3.17 (1.08-9.27)	-
≥1 preoperative areas or ≥2 density count[b]	-	-	2.80 (1.63-4.81)
C-statistic	0.64	0.68	0.68

CI, confidence interval; MHI-5, Mental Health Inventory; OR, odds ratio; Ref, reference; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index (categorized as 0, 1-19, and ≥20).
The following covariates were also included in the models but did not reach statistical significance: body mass index, Charlson Comorbidity Index, and Pain Catastrophizing Scale.
[a] Cumulative density count consisted of number of problematic musculoskeletal areas present preoperatively (either 0 or ≥1) and density count (either 0-1 or ≥2).
[b] In model 3 only, cumulative density count consisted of the following two groups: *1*) zero preoperative areas and 0 to 1 density count and *2*) one or more preoperative area or 2 or greater density count.

Table 7. Ordinal logistic regression models for 60-month WOMAC pain category by density count

Variable	Model 1 OR (95% CI)	Model 2 OR (95% CI)
Age per 5 y	0.99 (0.84-1.18)	0.98 (0.82-1.17)
Female sex	2.06 (1.17-3.63)	1.82 (1.02-3.23)
MHI-5 <68	3.92 (1.99-7.72)	3.31 (1.66-6.58)
Number of musculoskeletal problematic areas preoperatively, n (%)		
0	Ref	Ref
1	1.52 (0.76-3.07)	1.46 (0.72-2.97)
≥2	5.71 (2.19-14.87)	6.39 (2.44-16.73)
Density count		
0-1	-	Ref
2-4	-	2.09 (1.08-4.05)
≥5	-	3.53 (1.11-11.18)
C-statistic	0.68	0.71

CI, confidence interval; MHI-5, Mental Health Inventory; OR, odds ratio; Ref, reference; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index (categorized as 0, 1-19, and ≥20).
The following variables were included in the models but were not statistically significant: body mass index, Charlson Comorbidity Index, baseline WOMAC pain score, and Pain Catastrophizing Scale.

Sensitivity analysis: ordinal logistic regression by density count. When the number of preoperative problematic MSK areas and density count were incorporated as separate covariates in a model of WOMAC pain score at 60 months, the density count variable did not "step in" according to entry/exit criteria and needed to be forced into the model (model 2, Table 7). The adjusted OR was 2.09 (95% CI, 1.08-4.05) for a density count of 2 to 4 and 3.53 (95% CI, 1.11-11.18) for a density count of 5 or more in comparison with a density count of 0 to 1. Similarly, when the number of problematic MSK areas assessed preoperatively and density count were incorporated as separate covariates in a model of WOMAC function score at 60 months, the density count variable, once again, did not step into the model based on the entry/exit criteria. The adjusted OR was 1.86 (95% CI, 0.96-3.60) for a density count of 2 to 4 and 2.12 (95% CI, 0.65-6.85) for a density count of 5 or more in comparison with a density count of 0 to 1 (model 2, Appendix Table 8).

Table 8. Ordinal logistic regression models for 60-month WOMAC function category by density count

Variable	Model 1 OR (95% CI)	Model 2 OR (95% CI)
Age per 5 y	1.10 (0.93-1.30)	1.09 (0.92-1.29)
Female sex	1.60 (0.94-2.70)	1.43 (0.84-2.44)
MHI-5 <68	3.77 (1.88-7.55)	3.27 (1.62-6.62)
Number of musculoskeletal problematic areas preoperatively, n (%)		
0	Ref	Ref
1	1.62 (0.81-3.24)	1.55 (0.77-3.10)
≥2	8.30 (3.06-22.54)	8.97 (3.29-24.49)
Density count		
0-1	-	Ref
2-4	-	1.86 (0.96-3.60)
≥5	-	2.12 (0.65-6.85)
C-statistic	0.67	0.69

CI, confidence interval; MHI-5, Mental Health Inventory; OR, odds ratio; Ref, reference; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index (categorized as 0, 1-19, and ≥20).
The following variables were included in the models but were not statistically significant: body mass index, Charlson Comorbidity Index, baseline WOMAC function score, and Pain Catastrophizing Scale.

DISCUSSION

In this cohort of elective TKR recipients, the presence of problematic MSK areas before TKR and the development of new problematic MSK areas after surgery, as represented by cumulative density count, was significantly associated with worse 60-month WOMAC pain and function scores, even after adjustment for age, sex, baseline WOMAC function scores, and preoperative levels of anxiety and depression (MHI-5 <68). These findings highlight that both preoperative and postoperative problematic MSK complaints appear to be associated with chronic knee pain after TKR. This analysis also highlights that new areas of musculoskeletal pain arise frequently after TKR and influence long-term surgical outcomes. We are not able to ascertain the origin of these painful areas but highlight this as an important topic for future research.

Our results further support the idea that pain is a complex entity and difficult to predict. The addition of cumulative density count to the initial model led to rather modest increases in the C-statistic, underscoring that although this variable provides additional prediction, its marginal effect is not large. When we examined preoperative and postoperative problematic areas separately (Tables 7 and 8), these analyses also showed rather modest contributions to the prediction of pain and function scores.

Although no other studies have examined incident MSK complaints, our results align with previous literature on concomitant pain elsewhere and TKR outcomes. Wylde et al found that major depression and "number of pain problems elsewhere" were significant determinants of persistent pain 3 to 4 years after TKR (5). In that study, several years after surgery, participants were asked whether they suffered from pain in the preceding 4 weeks in any of the following five body areas: hips, ankles or feet, upper limbs, neck, and back. They were also asked whether they suffered from the following five pain conditions: migraine, irritable bowel syndrome, fibromyalgia, tinnitus, or chronic fatigue syndrome. Responses from both questions were combined into a single score categorized as 0, 1 to 2, 3 to 4, and 5 or more. Subjects with a score of at least 5 had 11.8-fold greater odds (95% CI, 5.33-26.07) of being grouped in a worse WOMAC pain category after TKR than a better one compared with those with a score of 0. In another study by Liu et al, the authors identified postoperative "presence of pain in other areas of the body" as a risk factor for persistent and severe postsurgical knee pain, which was defined as a pain score of at least 5/10, 1 year after TKR (OR, 1.54; 95% CI, 1.25-1.90) (20). Although neither of these studies explicitly measured incident MSK complaints after TKR, the findings document an association between multifocal pain and persistent postsurgical knee pain.

Indeed, multiple studies have demonstrated a link between widespread pain sensitization and chronic postsurgical knee pain. Whether a connection can be drawn between incident MSK areas and hyperalgesia is still unknown. Lundblad et al showed that low preoperative pain thresholds measured by electrical stimulation were significantly associated with worse visual analog scale (VAS) knee pain scores 18 months after TKR (21). Wylde et al found that lower pre-TKR pressure pain thresholds in the forearm (greater widespread pain sensitization) were correlated, albeit modestly, with more severe 1-year WOMAC pain scores (r = 0.37; P = 0.008) (22). Similarly, greater preoperative temporal summation of pain was correlated with worse VAS scores 12 months after TKR (r = 0.24; P = 0.037) (23). Although it is possible that incident MSK areas after TKR are a result of focal MSK processes, it is also possible that they reflect diffuse pain syndromes that predispose individuals to problematic MSK areas. Further research is needed to discern the exact mechanism.

These data may help clinicians set appropriate expectations with patients in discussions about TKR outcomes. Multiple studies have demonstrated that patient satisfaction with TKR is associated with patient expectations, which can be modified through preoperative patient education (24–28). Frank conversations between providers and patients regarding the risk of new sites of limitation and attendant risk of greater pain over follow-up may help minimize postoperative dissatisfaction. These concepts are important but also complex and will require careful explanation.

Although this study had a number of strengths, including large sample size, 5-year follow-up, and a comprehensive assessment of incident and persistent MSK pain, we note several limitations. The study cohort was derived from a single tertiary academic center, limiting generalizability, although the age range and sex breakdown of the cohort was representative of patients undergoing TKR in the United States (29). We did not ascertain laterality (left versus right) of problematic MSK areas (except for knees) or differentiate between certain joints (ie, hands/wrists/arms/shoulders and ankles/feet were grouped together). This may have led to underreporting of problematic joints. Subjects who did not provide data on specific problematic areas (either because they did not complete the questionnaire or did not complete specific items) were assumed to have no limitations in these areas. This is a conservative assumption and seems unlikely to influence results given that 80% of subjects completed all five questionnaires and 92% completed four or five questionnaires. Nevertheless, it is possible that the 15% of total responses that were missing and assumed to be nonproblematic could have led to miscalculation of ORs. It is difficult to know exactly how this missing data would have impacted the association between density count and 60-month outcomes. We also excluded subjects who did not complete the 60-month follow-up questionnaire. Excluded subjects had greater baseline anxiety and depression scores than included subjects. Based on our previous work, these participants would have likely had high density and cumulative density counts, leading to worse 60-month WOMAC scores (10). When calculating density count, we also gave equal weight to each area at each time point; for example, a single MSK problematic area present over three different time points had the same density count as three different MSK problematic areas present at one time point. Although it is possible these two scenarios impacted 60-month outcomes differently, we felt that combining number of problematic areas and time into

a single variable, the density count, was the most efficient method of capturing and evaluating the impact of these data. Finally, it is important to note that we did not evaluate change from baseline for longitudinally assessed instruments such as CCI and MHI-5. As such, it is possible baseline data for factors such as BMI, CCI, MHI-5, and PCS changed over time and that 60-month WOMAC scores could have been influenced in unmeasured ways.

Pre-existing and incident problematic MSK areas after TKR were associated with modest worsening in 60-month WOMAC pain and function scores. These findings underscore the complexity of predicting problematic and functional outcomes following TKR. Clinicians may find these associations useful in providing risk stratification and counseling both preoperatively and over the course of follow-up after TKR.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual contact, and all authors approved the final version to be published. Dr. Zhang had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Zhang, Selzer, Losina, Collins, Katz.
Acquisition of data. Selzer, Losina, Collins, Katz.
Analysis and interpretation of data. Zhang, Selzer, Losina, Collins, Katz.

REFERENCES

1. Deshpande BR, Katz JN, Solomon DH, Yelin EH, Hunter DJ, Messier SP, et al. Number of persons with problematic knee osteoarthritis in the us: impact of race and ethnicity, age, sex, and obesity. Arthritis Care Res 2016;68:1743–50.
2. Felson DT. Clinical practice. Osteoarthritis of the knee [published erratum appears in N Engl J Med 2006;354:2520]. N Engl J Med 2006;354:841–8.
3. McAlindon TE, Bannuru RR, Sullivan MC, Arden NK, Berenbaum F, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthritis Cartilage 2014;22:363–88.
4. Dowsey MM, Gunn J, Choong PF. Selecting those to refer for joint replacement: who will likely benefit and who will not? [Review]. Best Pract Res Clin Rheumatol 2014;28:157–71.
5. Wylde V, Hewlett S, Learmonth ID, Dieppe P. Persistent pain after joint replacement: prevalence, sensory qualities, and postoperative determinants. Pain 2011;152:566–72.
6. Brander VA, Stulberg SD, Adams AD, Harden RN, Bruehl S, Stanos SP, et al. Predicting total knee replacement pain: a prospective, observational study. Clin Orthop Relat Res 2003:27–36.
7. Kahn TL, Soheili AC, Schwarzkopf R. Poor WOMAC scores in contralateral knee negatively impact TKA outcomes: data from the osteoarthritis initiative. J Arthroplasty 2014;29:1580–5.
8. Perruccio AV, Power JD, Evans HM, Mahomed SR, Gandhi R, Mahomed NN, et al. Multiple joint involvement in total knee replacement for osteoarthritis: effects on patient-reported outcomes. Arthritis Care Res 2012;64:838–46.
9. Collados-Maestre I, Lizaur-Utrilla A, Martinez-Mendez D, Marco-Gomez L, Lopez-Prats FA. Concomitant low back pain impairs outcomes after primary total knee arthroplasty in patients over 65 years: a prospective, matched cohort study. Arch Orthop Trauma Surg 2016;136:1767–71.
10. Zhang M, Selzer F, Losina E, Collins JE, Katz JN. Musculoskeletal problematic areas after total knee replacement for osteoarthritis. ACR Open Rheumatol 2019;1:373–81.
11. Losina E, Collins JE, Daigle ME, Donnell-Fink LA, Prokopetz JJZ, Strnad D, et al. The AViKA (Adding Value in Knee Arthroplasty) postoperative care navigation trial: rationale and design features. BMC Musculoskelet Disord 2013;14:290.
12. Losina E, Collins JE, Wright J, Daigle ME, Donnell-Fink LA, Strnad D, et al. Postoperative care navigation for total knee arthroplasty patients: A randomized controlled trial. Arthritis Care Res 2016;68:1252–9.
13. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833–40.
14. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373–83.
15. Berwick DM, Murphy JM, Goldman PA, Ware JE, Barsky AJ, Weinstein MC. Performance of a five-item mental health screening test. Medical Care 1991;29:169–76.
16. Yamazaki S, Fukuhara S, Green J. Usefulness of five-item and three-item Mental Health Inventories to screen for depressive symptoms in the general population of Japan. Health Qual Life Outcomes 2005;3:48.
17. Rumpf HJ, Meyer C, Hapke U, John U. Screening for mental health: validity of the MHI-5 using DSM-IV Axis I psychiatric disorders as gold standard. Psychiatry Res 2001;105:243–53.
18. Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: development and validation. Psychological Assessment 1995;7:524–32.
19. Riddle DL, Wade JB, Jiranek WA, Kong X. Preoperative pain catastrophizing predicts pain outcome after knee arthroplasty. Clin Orthop Relat Res 2010;468:798–806.
20. Liu SS, Buvanendran A, Rathmell JP, Sawhney M, Bae JJ, Moric M, et al. A cross-sectional survey on prevalence and risk factors for persistent postsurgical pain 1 year after total hip and knee replacement. Reg Anesth Pain med 2012;37:415–22.
21. Lundblad H, Kreicbergs A, Jansson KA. Prediction of persistent pain after total knee replacement for osteoarthritis. J Bone Joint Surg Br 2008;90:166–71.
22. Wylde V, Palmer S, Learmonth ID, Dieppe P. The association between pre-operative pain sensitisation and chronic pain after knee replacement: an exploratory study. Osteoarthritis Cartilage 2013;21:1253–6.
23. Petersen KK, Arendt-Nielsen L, Simonsen O, Wilder-Smith O, Laursen MB. Presurgical assessment of temporal summation of pain predicts the development of chronic postoperative pain 12 months after total knee replacement. Pain 2015;156:55–61.
24. Noble PC, Conditt MA, Cook KF, Mathis KB. The John Insall Award: patient expectations affect satisfaction with total knee arthroplasty. Clin Orthop Relat Res 2006;452:35–43.
25. Scott CE, Howie CR, MacDonald D, Biant LC. Predicting dissatisfaction following total knee replacement: a prospective study of 1217 patients. J Bone Joint Surg Br 2010;92:1253–8.
26. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KD. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? [full-length article]. Clinical Orthop Relat Res 2010;468:57–63.
27. Neuprez A, Delcour JP, Fatemi F, Gillet P, Crielaard JM, Bruyere O, et al. Patients' expectations impact their satisfaction following total hip or knee arthroplasty. PLoS One 2016;11:e0167911.
28. Mancuso CA, Graziano S, Briskie LM, Peterson MG, Pellicci PM, Salvati EA, et al. Randomized trials to modify patients' preoperative expectations of hip and knee arthroplasties. Clinical Orthop Relat Res 2008;466:424–31.
29. Healthcare Cost and Utilization Project: Nationwide Inpatient Sample (NIS). Agency for Healthcare Research and Quality. 2014 URL: https://hcupnet.ahrq.gov/