Am J Transl Res 2026;18(4):3360-3371
www.ajtr.org /ISSN:1943-8141/AJTR0171614

Original Article

Development and validation of a clinically available risk model for medium-to-high opioid consumption after total knee arthroplasty

Yonghong Ding[1,2], Kuilin Wu[2,3], Zongwang Zhang[1,2]

[1]Department of Anesthesiology, Liaocheng People's Hospital, Shandong University, Jinan, Shandong, China; [2]Department of Anesthesiology, Liaocheng People's Hospital, Liaocheng, Shandong, China; [3]Shandong First Medical University, Jinan, Shandong, China

Received January 22, 2026; Accepted March 26, 2026; Epub April 15, 2026; Published April 30, 2026

Abstract: Objectives: To develop and internally validate a clinically interpretable preoperative model for predicting medium-to-high opioid consumption within 72 h after unilateral total knee arthroplasty (TKA). Methods: We retrospectively analyzed 806 patients who underwent primary unilateral TKA between October 2017 and September 2021. Patients were randomly divided into a training cohort (70%) and a validation cohort (30%). The primary outcome was total opioid consumption within the first 72 postoperative hours, categorized as low versus medium-to-high consumption. Baseline and perioperative variables were first compared using appropriate univariable statistical tests, and variables showing potential associations, together with clinically relevant factors, were entered into a multivariable logistic regression model with bidirectional stepwise selection. The final model was presented as a nomogram. Discrimination, calibration, and clinical utility were assessed using the area under the receiver operating characteristic curve (AUC), calibration plots with intercepts and slopes, Hosmer-Lemeshow testing, and decision curve analysis. Results: Female sex, diabetes mellitus, higher body mass index, and higher preoperative day-1 serum interleukin-6 (IL-6) levels were independent predictors of medium-to-high postoperative opioid consumption. The model showed acceptable discrimination in the training cohort (AUC 0.727) and modest discrimination in the validation cohort (AUC 0.665). Calibration was acceptable in both cohorts (training: intercept 0.012, slope 0.953; validation: intercept 0.009, slope 0.941). Decision curve analysis supported potential clinical utility across relevant threshold probabilities. Conclusions: A simple preoperative model based on routinely available demographic, clinical, and inflammatory variables may help identify patients at increased risk of higher opioid requirements after unilateral TKA and may support individualized multimodal analgesic planning.

Keywords: Total knee arthroplasty, opioid consumption, risk prediction, interleukin-6, body mass index, postoperative pain

Introduction

Total knee arthroplasty (TKA) is a widely performed intervention for end-stage knee osteoarthritis, and its use continues to increase as populations age [1]. Although TKA reliably improves pain and function, a substantial proportion of patients still experience considerable acute pain, particularly during the first 72 hours after surgery [2]. Inadequate analgesia may delay mobilization and rehabilitation, prolong hospitalization, and increase the risk of persistent postsurgical pain [3]. Conversely, excessive opioid exposure is associated with adverse effects such as nausea, respiratory depression, delirium, and prolonged postoperative opioid use [4-6]. These considerations support a multimodal, opioid-sparing analgesic approach; however, considerable interindividual variability in postoperative opioid requirements remains even under standardized perioperative care [7].

A wide-ranging literature describes demographic, psychosocial, and clinical factors associated with postoperative opioid use following orthopedic procedures, including younger age, female sex, prior opioid exposure, anxiety, depression, and several comorbidities [8-10]. Mechanistic studies additionally implicate sys-

https://doi.org/10.62347/SSCM5032

temic inflammation as a major driver of postoperative pain with interleukin-6 (IL-6) most prominently correlating with tissue injury and pain sensitivity [11]. Peri-operative glucocorticoids, including dexamethasone, decrease inflammation and are associated with lower pain scores and opioid consumption in TKA and other surgeries [12-15]. However, most such prior studies examined average treatment effects or long term trajectories of opioid usage, not who will require a higher opioid dose in the early postoperative setting [16].

More recently, non-deep learning statistical and machine-learning approaches have evaluated whether opioid consumption or prolonged use after surgery can be predicted [17-19]. Only a handful of groups have investigated opioid requirements out to 72 hours after TKA - just when pain and opioid requirements peak and rehabilitation begins [20, 21]. Existing predictive models are often based on complex algorithms or perioperative variables that are not readily obtainable or interpretable for common preoperative decisions, limiting their viability in routine clinical practice [22]. These gaps in predictive modeling highlight the need for a simple and clinically interpretable preoperative tool to stratify patients based on their medium-to-high likelihood of postoperative opioid consumption.

Recognizing this gap, we sought to develop and validate a preoperative risk model for medium-to-high postoperative opioid consumption following unilateral TKA. Using only variables available before surgery, including demographic factors, comorbidities such as diabetes mellitus, body mass index, and preoperative day-1 serum IL-6 level [23], we constructed a multivariable prediction model and translated it into a bedside nomogram. The model demonstrated acceptable discrimination, calibration, and potential clinical utility for individualized perioperative analgesic planning [24].

Methods

Patients and study design

We did a retrospective observational study on an original cohort of patients who underwent TKA at our hospital between October 2017 and September 2021. In view of the retrospective nature of this study, the necessary patient information was retrieved from the medical records database of Liaocheng People's Hospital. This allowed us to extract demographic data (e.g., age, sex), procedures, prescriptions, inpatient or outpatient status, and laboratory data.

Inclusion criteria were male and female patients over the age of 18 years who underwent primary TKA during the study period for knee OA, and were classified as ASA I-III. Exclusion criteria were: 1) joint re-replacement surgery during the study period; 2) rheumatoid arthritis; 3) malignancy/chronic heart failure/renal disorder diagnosed prior to TKA; 4) diagnosis of infection during hospitalization; 5) incomplete serum IL-6 measurements; 6) missing BPI scores or unknown details of opioid consumption. On this basis, 62 patients were excluded from further analysis.

All TKAs were performed unilaterally under general anesthesia by a single group of board-certified orthopedic surgeons using consistent institutional protocols. Anesthesia was administered based on weight guidelines, and local anesthesia was not used in any cases. This study received special approval from the Institutional Review Board of Liaocheng People's Hospital (Approval No. 2018063). As this study is retrospective, the requirement for informed consent was waived.

For prediction model development, patients were randomly assigned to a training cohort (70%) and a validation cohort (30%) using a computer-generated random sequence with simple randomization. Baseline characteristics, including age, sex, BMI, comorbidities, ASA class, psychological status, and IL-6 levels, were well balanced between the two cohorts, providing a robust basis for model development and validation.

Data collection

Data were obtained from the electronic patient record, anesthesia information system, and inpatient medical record system. Candidate preoperative or perioperative variables considered during model development included age, sex, BMI, ASA class, smoking history, alcohol history, years of education, hypertension, diabetes mellitus, preoperative serum IL-6, preoperative pain-related information, preoperative

analgesic use, intraoperative NSAID use, operative duration, and tourniquet duration. Because preoperative serum IL-6 showed a skewed distribution, it was summarized as median (interquartile range) and analyzed using nonparametric methods where appropriate.

Venous blood samples were collected between 08:00 and 10:00 on the day before surgery and processed according to standard institutional procedures. Serum IL-6 concentration was measured in the central laboratory of Liaocheng People's Hospital, and the results were expressed in pg/mL. Routine internal quality-control procedures were performed throughout the study period to ensure analytical reliability.

For each patient, demographic data, medical history, anesthesia and surgery details, and follow-up information were obtained from clinical databases and medical records. Demographic data included age, sex, BMI, and ASA physical status. Medical history included hypertension, diabetes mellitus, smoking history, alcohol history, and years of education. Candidate perioperative variables also included preoperative analgesic use, intraoperative NSAID use, operative duration, and tourniquet duration; these variables were considered during model development but were not retained in the final model because they did not meet the prespecified selection criteria or did not materially improve model performance. Prior long-term opioid exposure was not consistently available in structured records and therefore could not be analyzed directly.

In accordance with our multimodal postoperative pain regimen, patients received an ultrasound-guided adductor canal (subsartorial saphenous nerve) block and patient-controlled intravenous analgesia (PCIA) until the afternoon of postoperative day 3. Pain scores were recorded using the Brief Pain Inventory (BPI) on postoperative days 1, 2, and 3.

Primary outcome

The primary outcome was total opioid consumption in the postoperative period for the first 72-hours after unilateral TKA. Total opioid use (sufentanil equivalents), derived from electronically monitored doses of a patient-controlled analgesia infusion of sufentanil, was interpreted according to distributional information and clinical relevance as low versus medium-to-high consumption and used as binary outcome in models developed. The secondary outcome was pain intensity (BPI on post-op days 1-3); pain scores were not included in the final model as a predictor, since they are unavailable pre-operatively. Postoperative BPI scores on postoperative days 1-3 were collected for descriptive purposes only and were not analyzed using a repeated-measures inferential model, because pain intensity was not part of the primary preoperative prediction analysis.

Sample size

Sample size adequacy was determined using the "events per variable" (EPV) rule for prediction modeling. With four final multivariable predictors and > 400 medium-to-high-consumption events in the full cohort, the minimum required (≥ 10 per variable) EPV was exceeded, which allows reliable estimation of coefficients and stable model performance.

Data management and statistical analysis

All analyses adhered to the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) guidelines. Continuous variables were summarized as mean ± standard deviation or median (IQR) and compared using Student's t test or Wilcoxon rank-sum test according to distributional characteristics; categorical variables were compared using the chi-square test or Fisher's exact test, as appropriate.

Baseline and perioperative variables were first compared between groups using appropriate univariable statistical tests. Variables showing potential associations in univariable comparisons, together with clinically relevant factors, were then entered into a multivariable logistic regression model to identify independent predictors of medium-to-high postoperative opioid consumption. Bidirectional stepwise selection was applied during multivariable model development, with $P < 0.05$ for entry and $P > 0.10$ for removal.

The final model identified sex, diabetes mellitus, body mass index, and preoperative day-1 serum IL-6 level as independent predictors of

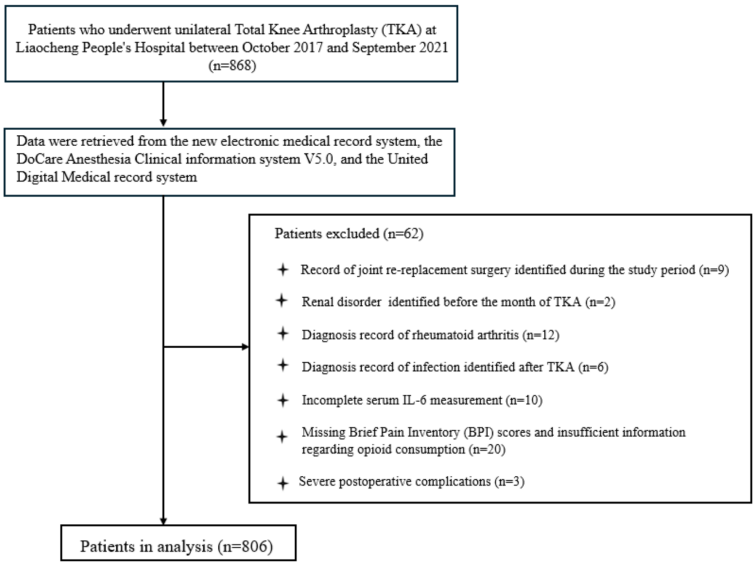

Figure 1. The flowchart of patient selection.

medium-to-high opioid use. A nomogram was then constructed on the basis of the regression coefficients to facilitate bedside application.

Predictive performance was evaluated in both cohorts using receiver operating characteristic (ROC) curves and areas under the curve (AUCs). Calibration was assessed using calibration plots, the Hosmer-Lemeshow test, and calibration intercepts and slopes. Well-calibrated models have an intercept close to 0 and a slope close to 1. Decision curve analysis was used to quantify the net clinical benefit of model-guided decision-making across threshold probabilities.

Results

Patient selection and cohort derivation

A total of 868 patients who underwent primary unilateral total knee arthroplasty (TKA) between October 2017 and September 2021 were screened. After exclusion of 62 patients because of prior joint revision during the study period (n=9), recent renal impairment (n=2), rheumatoid arthritis (n=12), postoperative infection (n=6), incomplete serum IL-6 measurements (n=10), missing Brief Pain Inventory (BPI) scores or undocumented opioid consumption (n=20), and severe postoperative complications (n=3), 806 patients were included in the final analysis. The cohort was randomly divided into a training set (n=565) and a validation set (n=241) (**Figure 1**).

Baseline characteristics

Baseline demographic and clinical characteristics were comparable between the training and validation cohorts. There were no significant between-cohort differences in age, sex, American Society of Anesthesiologists (ASA) class, body mass index (BMI), smoking history, alcohol history, years of education, Hospital Anxiety and Depression Scale-anxiety (HADS-A) score, Hospital Anxiety and Depression Scale-depression (HADS-D) score, coronary heart disease, hypertension, diabetes mellitus, or preoperative day-1 serum IL-6 level (all $P > 0.05$). These findings indicated good baseline comparability between the 2 cohorts (**Table 1**).

Univariable analysis in the training cohort

In the training cohort, patients were stratified into low- and medium-to-high-consumption groups based on postoperative opioid use. Univariable analyses showed that the medium-to-high-consumption group was younger and had a higher proportion of female patients than the low-to-medium-consumption group. In addition, smoking status, alcohol use, diabetes mellitus, and preoperative IL-6 level differed significantly between groups (all $P < 0.05$). Although the median IL-6 value was 6.0 pg/ml in both groups, the distribution of IL-6 differed between groups, with a higher interquartile range in the medium-to-high-consumption group (6.0-8.0 pg/ml vs 5.0-7.0 pg/ml), and the rank-sum test showed a significant difference ($Z=-4.710$, $P < 0.001$). By contrast, ASA class, BMI, years of education, HADS-A score, HADS-D score, coronary heart disease, hypertension, preoperative WOMAC score, preoperative pain score, preoperative analgesic use, intraoperative NSAID use, operative duration, and tourniquet duration were not significantly associated with opioid-consumption grouping (all $P > 0.05$). Variables showing potential associations in uni-

Table 1. Characteristics of Patients in the Training and Validation Cohorts

Characteristics	Total	Training Set	Validation Set	Test statistic	P
Age, (mean ± SD), years	66.5±6.6	66.5±6.6	66.4±6.9	*t*=0.252	0.801†
Sex, n (%)				χ^2=0.018	0.892
Male	385 (47.77)	269 (47.61)	116 (48.13)		
Female	421 (52.23)	296 (52.39)	125 (51.87)		
ASA (%)				χ^2=0.114	0.945
I	56 (6.95)	40 (7.08)	16 (6.64)		
II	683 (84.74)	479 (84.78)	204 (84.65)		
III	67 (8.31)	46 (8.14)	21 (8.71)		
BMI, (mean ± SD)	26.9±3.1	26.8±3.0	26.9±3.2	*t*=0.373	0.710†
Smoking (%)				χ^2=0.426	0.514
yes	237 (29.4)	170 (30.09)	67 (27.80)		
no	569 (70.6)	395 (69.91)	174 (72.2)		
Drinking (%)				χ^2=3.268	0.071
yes	333 (41.32)	245 (43.36)	88 (36.51)		
no	473 (58.68)	320 (56.64)	153 (63.49)		
Education years, Median (IQR), years	9.00 (6.00, 12.00)	9.00 (6.00, 13.00)	9.00 (5.00, 12.00)	*Z*=-1.832	0.067#
HADS-A, Median (IQR)	1.00 (0.00, 2.00)	1.00 (0.00, 2.00)	1.00 (0.00, 2.00)	*Z*=-0.681	0.495#
HADS-D, Median (IQR)	1.00 (0.00, 3.00)	1.00 (0.00, 3.00)	1.00 (0.00, 3.00)	*Z*=-0.362	0.717#
Coronary heart disease (%)				χ^2=0.118	0.731
yes	78 (9.68)	56 (9.91)	22 (9.13)		
no	728 (90.32)	509 (90.09)	219 (90.87)		
Hypertension (%)				χ^2=0.028	0.868
yes	381 (47.27)	266 (47.08)	115 (47.72)		
no	425 (52.73)	299 (52.92)	126 (52.28)		
Diabetes mellitus (%)				χ^2=0.498	0.480
yes	180 (22.33)	130 (23.01)	50 (20.75)		
no	626 (77.67)	435 (76.99)	191 (79.25)		
preoperative day-1 serum IL-6 level, Median (IQR), (pg/ml)	6.00 (5.00, 8.00)	6.00 (5.00, 8.0)	6.00 (5.00, 8.00)	*Z*=-0.357	0.721#

Data are presented as mean ± standard deviation for continuous variables and n (%) for categorical variables. ASA, American Society of Anesthesiologists; BMI, body mass index. "#" means rank sum test; "†" means t-test.

variable comparisons were carried forward for multivariable model development, together with clinically relevant factors (**Table 2**).

Multivariable model development

Variables identified in the univariable analysis were entered into the multivariable logistic regression model. Four variables remained independently associated with medium-to-high opioid consumption within 72 h after TKA. Female sex was associated with higher odds of the outcome than male sex (OR 3.296, 95% CI 2.159-5.032; P < 0.001). Diabetes mellitus was also independently associated with the outcome (OR 2.096, 95% CI 1.214-3.621; P=0.008). Higher serum IL-6 level on the day before surgery was positively associated with the outcome, with an OR of 1.296 per pg/mL increase (95% CI 1.155-1.454; P < 0.001). Higher BMI was likewise independently associated with the outcome, with an OR of 1.083 per kg/m^2 increase (95% CI 1.013-1.159; P=0.020) (**Table 3**).

Nomogram construction

A nomogram was constructed based on the final multivariable logistic regression model to estimate the probability of medium-to-high opioid consumption within 72 h after TKA. The nomogram incorporated sex, diabetes mellitus, preoperative day-1 serum IL-6 level, and BMI. For score assignment, each predictor was weighted according to its standardized regression coefficient in the final model, and the predictor with the largest absolute coefficient was scaled to 100 points. The scores assigned to the remaining predictors were calculated proportionally. The total score, obtained by summing the points for all predictors, was then mapped to the corresponding predicted probability of the outcome (**Figure 2**).

Table 2. Baseline and perioperative variables in patients with low-to-medium versus medium-to-high postoperative opioid consumption

Variables	Total	Low consumption	Medium-to-high consumption	Test statistic	P
Age, (mean ± SD), years	66.5±6.6	67.7±6.1	66.2±6.7	t=2.484	0.013†
Sex, n (%)				χ^2=34.020	< 0.001
Male	269 (47.61)	96 (69.06)	173 (40.61)		
Female	296 (52.39)	43 (30.94)	253 (59.39)		
ASA (%)				χ^2=0.977	0.613
I	40 (7.08)	9 (6.47)	31 (7.28)		
II	479 (84.78)	116 (83.45)	363 (85.21)		
III	46 (8.14)	14 (10.07)	32 (7.51)		
BMI, (mean ± SD)	26.8±3.0	26.4±2.9	27.0±3.1	t=1.937	0.053†
Smoking (%)				χ^2=4.698	0.030
yes	170 (30.09)	52 (37.41)	118 (27.70)		
no	395 (69.91)	87 (62.59)	308 (72.30)		
Drinking (%)				χ^2=6.292	0.012
yes	245 (43.36)	73 (52.52)	172 (40.38)		
no	320 (56.64)	66 (47.48)	254 (59.62)		
Coronary heart disease (%)				χ^2=0.005	0.942
yes	56 (9.91)	14 (10.07)	42 (9.86)		
no	509 (90.09)	125 (89.93)	384 (90.14)		
Hypertension (%)				χ^2=3.500	0.061
yes	266 (47.08)	75 (53.96)	191 (44.84)		
no	299 (52.92)	64 (46.04)	235 (55.16)		
Diabetes mellitus (%)				χ^2=7.733	0.005
yes	130 (23.01)	20 (14.39)	110 (25.82)		
no	435 (76.99)	119 (85.61)	316 (74.18)		
Education years, Median (IQR), years	9.0 (6.0, 13.0)	9.0 (5.0, 12.0)	9.0 (6.0, 13.0)	Z=-0.484	0.628#
HADS-A, Median (IQR)	1.0 (0.0, 2.0)	2.0 (0.0, 3.0)	1.0 (0.0, 2.0)	Z=1.085	0.277#
HADS-D, Median (IQR)	1.0 (0.0, 3.0)	1.0 (0.0, 3.0)	1.0 (0.0, 3.0)	Z=-0.950	0.342#
preoperative day-1 serum IL-6 level, Median (IQR), (pg/ml)	6.0 (5.0, 8.0)	6.0 (5.0, 7.0)	6.0 (6.0, 8.0)	Z=-4.710	< 0.001#
Preoperative WOMAC score, (mean ± SD)	43.7±16.8	41.8±17.6	44.3±16.5	t=1.528	0.127†
Preoperative pain score, Median (IQR)	3.0 (2.0, 5.0)	3.0 (2.0, 5.0)	3.0 (2.0, 5.0)	Z=-1.114	0.265#
Preoperative analgesic use (%)				χ^2=2.642	0.104
yes	182 (32.21)	37 (26.62)	145 (34.04)		
no	383 (67.79)	102 (73.38)	281 (65.96)		
Surgery time, Median (IQR), min	112.0 (104.0, 119.0)	112.0 (104.0, 118.0)	112.0 (103.0, 120.0)	Z=-0.780	0.435#
Tourniquet time, (mean ± SD), min	68.4±11.4	69.5±11.9	68.0±11.3	t=1.371	0.171†

Data are presented as mean ± standard deviation for continuous variables and n (%) for categorical variables. ASA, American Society of Anesthesiologists; BMI, body mass index; Non-Steroidal Anti-Inflammatory Drugs. "#" means rank sum test; "†" means t-test.

Model discrimination, calibration, and clinical utility

The model showed acceptable discrimination in the training cohort, with an area under the receiver operating characteristic curve (AUC) of 0.727 (95% CI 0.680-0.775), and modest discrimination in the validation cohort, with an AUC of 0.665 (95% CI 0.585-0.744) (**Figure 3**). Calibration analysis demonstrated acceptable agreement between predicted and observed probabilities. In the training cohort, the calibration intercept and slope were 0.012 and 0.953, respectively; in the validation cohort, they were 0.009 and 0.941, indicating adequate calibration in both cohorts. Hosmer-Lemeshow testing did not indicate significant miscalibration in either cohort (training cohort: P=0.3512; validation cohort: P=0.6040) (**Figure 4**). Decision curve analysis supported the clinical utility of the model (**Figure 5**). In both cohorts, the nomogram provided a greater net benefit than the treat-all and treat-none strategies across a broad range of clinically relevant threshold probabilities, with the advantage being more evident at relatively higher thresholds.

Table 3. Multivariate Analysis of the Training Cohort

Variables	Estimate	Standard error	Wald	P	OR (95% CI)
Intercept	-3.372	1.028	10.751	0.001	
Sex, n (%)					
Male					Ref (1.000)
Female	1.193	0.216	30.541	< 0.001	3.296 (2.159, 5.032)
Diabetes mellitus (%)					
no					Ref (1.000)
yes	0.740	0.279	7.047	0.008	2.096 (1.214, 3.621)
Preoperative day-1 serum IL-6 level (per 1 pg/mL)	0.259	0.059	19.367	< 0.001	1.296 (1.155, 1.454)
BMI	0.080	0.034	5.411	0.020	1.083 (1.013, 1.159)

Abbreviations: OR, odds ratio; CI, confidence interval; BMI, body mass index; IL-6, interleukin-6. Male sex and no diabetes were the reference categories. IL-6 was analyzed per 1 pg/mL increase.

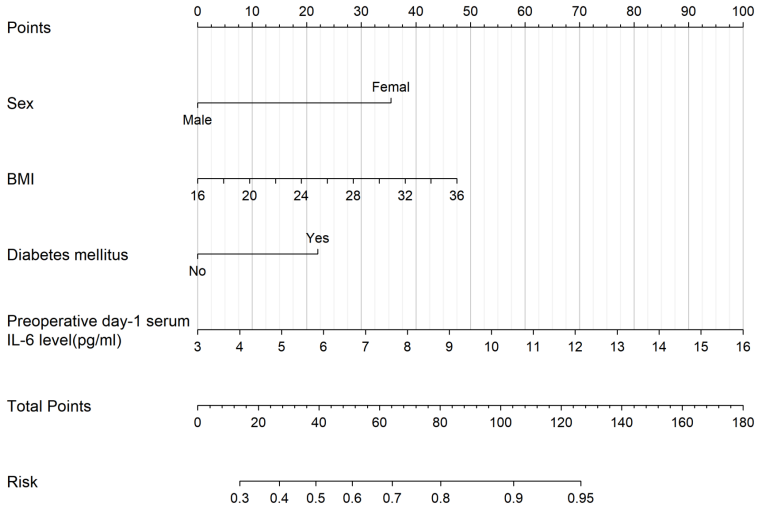

Figure 2. Nomogram predicting medium-to-high opioid consumption within 72 hours after Total Knee Arthroplasty (TKA). Nomogram estimating mediumtohigh opioid use within 72 h after TKA. Points are allocated for sex, diabetes status, BMI and preoperative day1 IL6 concentration; their sum yields a total score, which maps to the predicted risk on the "Risk" axis.

Discussion

The present study demonstrates that several preoperative demographic, clinical, and inflammatory features can help identify patients who are likely to require medium-to-high opioid doses during the first 72 hours after unilateral TKA. We developed a multivariable logistic model incorporating sex, diabetes status, BMI, and preoperative day-1 serum IL-6 level, and translated it into a clinically interpretable nomogram capable of estimating individualized postoperative opioid requirements. The performance of this model suggests that meaningful early risk stratification is possible using variables routinely available in standard perioperative care.

This study provides a clinically relevant framework for understanding interindividual variability in postoperative opioid requirements following TKA. Our results identified female sex as an independent predictor of higher postoperative opioid consumption after TKA. This finding is consistent with previous studies reporting sex-related differences in pain sensitivity, central sensitization, and opioid pharmacodynamics, which may partly explain the greater postoperative opioid requirements observed in female patients [25, 26]. These findings have also been documented in orthopedic, and general surgical experimental environments; thus it is clear that the impact of gender on analgesic therapies needs to be considered when developing personalized analgesic protocols. Additionally, the association between diabetes and opioid use supports previous research that indicates that diabetic individuals have altered mechanisms for transmitting pain signals, increased levels of systemic inflammation, and a decreased ability to return to normal function shortly after joint arthroplasty [27, 28]. By identifying diabetes as a definitive risk factor for the use of opioids within the first few weeks after surgery, this study reinforces the importance of considering diabetes in preoperative planning and education for all patients undergoing TKA.

Notably, preoperative day-1 serum IL-6 level concentration also emerged as a risk factor for

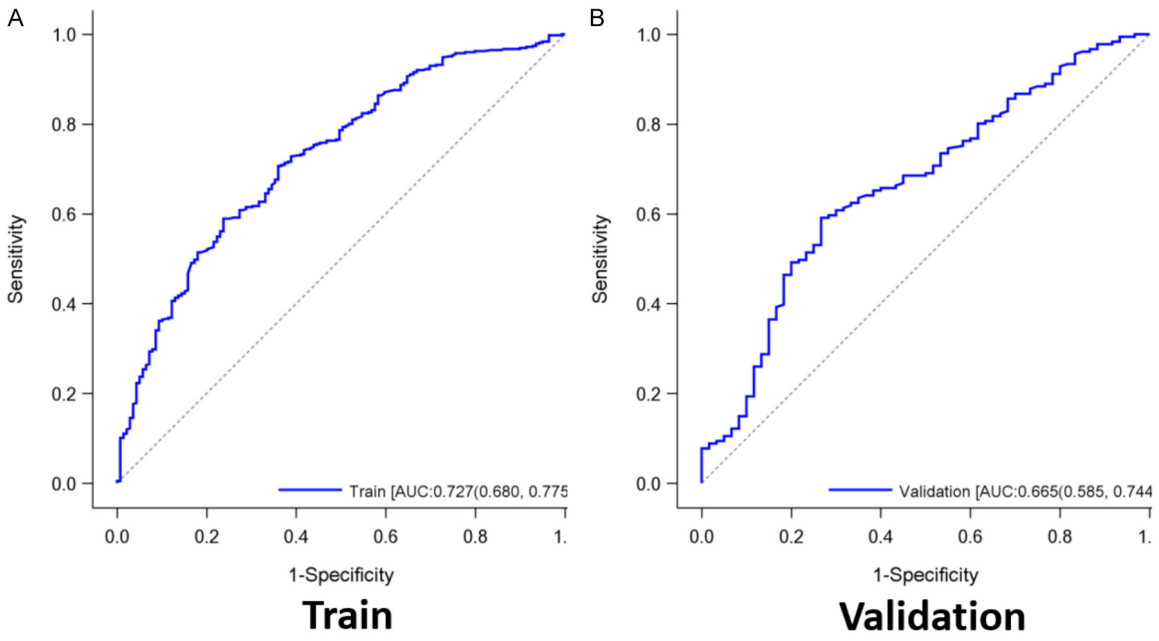

Figure 3. Receiver operating characteristic curves of the prediction model in the training and validation cohorts. ROC curves demonstrated moderate discrimination (AUC 0.727 and 0.665).

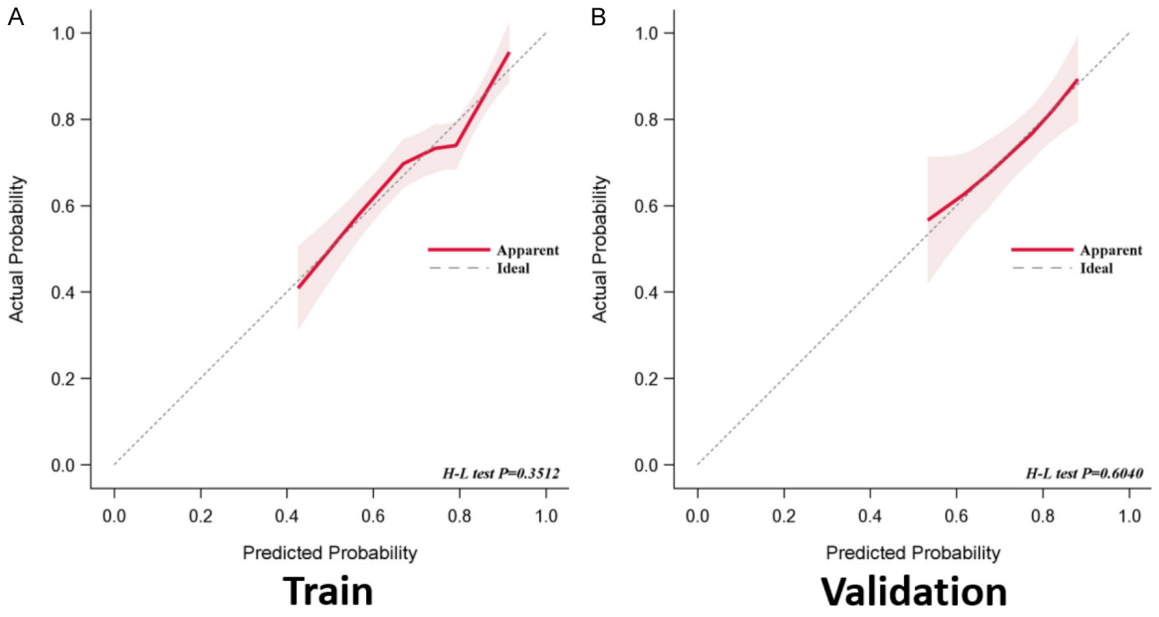

Figure 4. Calibration plots of the prediction model in the training and validation cohorts. Calibration curves showed good alignment between predicted and observed risks, with nonsignificant Hosmer-Lemeshow tests.

postoperative opioid requirement. IL-6 is one of the key cytokines signaling inflammatory pain, and elevated IL-6 levels are associated with increased postoperative pain severity, slower rehabilitation, and increased opioid use [14-16, 29]. Including IL-6 in our model provides a biologically meaningful operationalization of this mechanistic association. The operationalization of biological allostatic load factors (such as inflammatory biomarkers) contributing to pain and opioid use represents an advance over previous models based on mainly psychosocial or demographic factors, without incorporating inflammatory markers [12, 30]. The reported

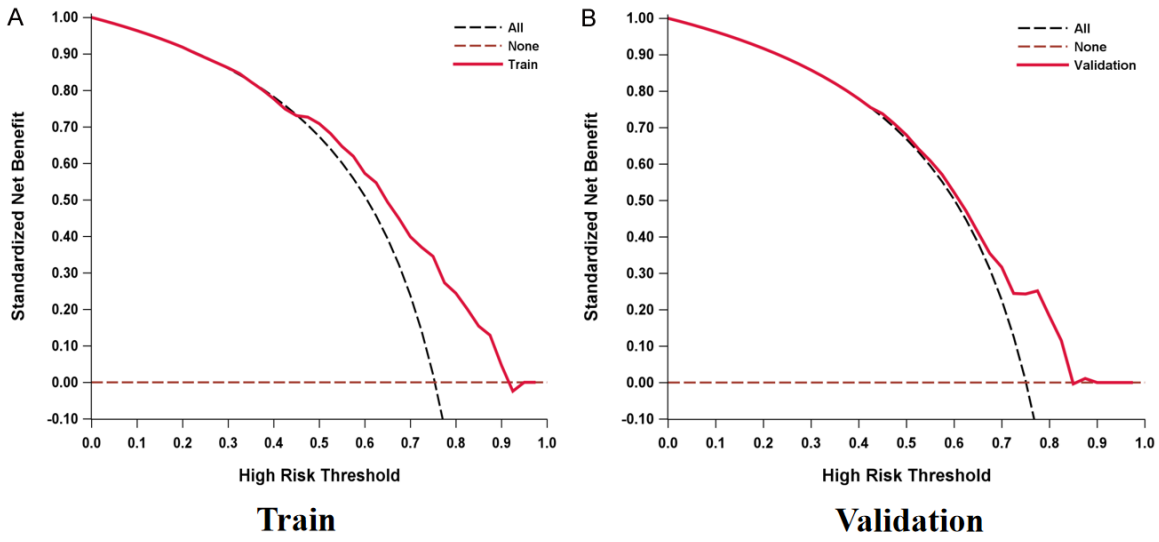

Figure 5. Decision curve analysis of the prediction model in the training and validation cohorts. Decision curve analysis indicated greater net clinical benefit than treat-all or treat-none strategies at threshold probabilities above 0.60.

odds ratio reflects the effect of each 1-unit increase in IL-6 level, suggesting that even modest elevations may indicate clinically meaningful inflammatory activity, particularly when considered together with demographic and clinical risk factors. Usage of IL-6 as a continuous biomarker for risk stratification rather than a dichotomous variable appears to be justified. Although our model used only preoperative day-1 IL-6, accumulating mechanistic evidence supports a causal role of IL-6 signaling in amplifying nociceptive processing [31-35]. IL-6 binds IL-6R and the signal-transducing subunit gp130, activating downstream JAK2/STAT3 signaling in the spinal cord and glial cells, which can promote neuroinflammation, synaptic plasticity, and central sensitization - processes that plausibly increase acute postoperative pain burden and opioid demand [32-35]. Experimental work has shown that spinal JAK/STAT3 activation contributes to pain hypersensitivity and that blockade of this pathway can attenuate allodynia/hyperalgesia, supporting biological plausibility for incorporating IL-6 as a risk marker in perioperative prediction [35, 36]. However, postoperative IL-6 kinetics may better reflect the inflammatory response to surgical injury. Because serial postoperative IL-6 measurements were not routinely available in this retrospective cohort, we were unable to directly examine the association between postoperative IL-6 dynamics and 72-hour opioid consumption. This should be addressed in future prospective studies with standardized perioperative cytokine monitoring.

Although the present model demonstrated acceptable discrimination and calibration, we did not perform a formal head-to-head comparison with previously published models for predicting opioid consumption or prolonged opioid use after TKA, including machine-learning-based approaches [36-38]. Existing studies have reported higher discrimination in some settings; for example, machine-learning or penalized regression models for extended, persistent, or long-term opioid use after TKA or lower-extremity arthroplasty have reported AUCs ranging from approximately 0.79 to 0.96 [36, 37]. However, these studies differed substantially from ours in target outcomes, prediction windows, predictor sets, and study populations, which limits direct comparison of AUC, sensitivity, and specificity [38]. In particular, some prior models incorporated postoperative variables, such as postoperative day-1 opioid use, discharge prescriptions, or early post-discharge information, whereas our nomogram was intentionally designed as a clinically interpretable preoperative tool based on routinely available variables [37]. Importantly, we also did not quantify the incremental predictive value of this model over conventional clinical assessment. Therefore, the extent to which the nomogram improves risk identification beyond routine clinician judgment remains uncertain. Future stu-

dies should directly compare this model with existing regression- and machine-learning-based tools in external datasets and evaluate added value using reclassification and decision-curve methods in addition to conventional discrimination metrics [38, 39].

Several factors previously reported to influence postoperative opioid consumption, such as preoperative pain scores, preoperative opioid use, surgical duration, and tourniquet time, did not emerge as independent predictors in our final model. This may be explained by several reasons. First, in our cohort, preoperative pain scores and preoperative analgesic use were relatively homogeneous due to standardized preoperative management, potentially limiting their discriminative ability. Second, the lack of consistent documentation of preoperative opioid use in our retrospective database prevented its reliable assessment, which may have led to residual confounding. Third, operative and tourniquet durations were largely uniform under the standardized surgical protocols at our institution, reducing their variability and thus their predictive value. These factors underscore the importance of context-specific model development and highlight that predictors may differ across populations and practice settings. Future studies should explore these variables in more heterogeneous cohorts or with prospective data collection.

This study has several limitations. First, although internal validation was performed by randomly splitting the cohort, the model was not externally validated in independent datasets from other institutions or time periods, which limits generalizability. Second, the study was retrospective and based on a single center, so residual confounding cannot be fully excluded. Third, opioid consumption during the first 72 postoperative hours was used as the primary endpoint; therefore, the model does not address longer-term opioid use or persistent postsurgical pain. Fourth, although IL-6 provided biologically relevant inflammatory information, postoperative pain is influenced by multiple biological, psychological, and social factors that were not fully captured in the present dataset. Finally, stepwise regression may introduce model instability, and future studies should consider external validation and penalized regression approaches. The impact of this model on real-world clinical decision-making should be evaluated prospectively in pragmatic or randomized trials.

Future work should include multicenter external validation to assess generalizability to different clinical settings, as well as an extension of this model to additional measures of biological and patient-reported pain phenotyping for improved prediction. Embedding the risk model into electronic health records could facilitate automated risk stratification and decision support at the point of care. Ultimately, prospective interventional trials using risk stratified analgesic protocols based on preoperative predicted risk will be needed to determine if we can use these strategies to reduce postoperative opioid exposure and improve pain control and recovery after TKA.

Conclusion

In the present study, we found that female sex, diabetes mellitus, higher BMI, and elevated serum IL-6 levels on preoperative day 1 were independent predictors of early postoperative opioid requirements following unilateral TKA. Using these four factors, we created and validated a streamlined and clinically useful nomogram to estimate individualized postoperative opioid consumption risk. By allowing for early identification of patients at high risk for opioid use, the nomogram may help facilitate more refined multimodal analgesic approaches and reduced exposure to opioids. Additional multicenter studies are needed to confirm the generalizability of the nomogram as well as its clinical utility and relevance.

Disclosure of conflict of interest

None.

Address correspondence to: Zongwang Zhang, Department of Anesthesiology, Liaocheng People's Hospital, Shandong University, Jinan, Shandong, China. Tel +86-13346256809; E-mail zongwang0122@126.com

References

[1] van Deventer L, Bronstone A, Leonardi C, Bennett M, Yager P and Dasa V. A modern multimodal pain protocol eliminates the need for opioids for most patients following total knee arthroplasty: results from a retrospective comparative cohort study. J Exp Orthop 2023; 10: 20.

[2] Lacko M, Matuska M, Folvarsky M, Gharaibeh A, Lackova A, Polan P and Zamborsky R. A multimodal opioid-sparing pain management following total knee replacement. Bratisl Lek Listy 2022; 123: 449-454.
[3] Kehlet H, Jensen TS and Woolf CJ. Persistent postsurgical pain: risk factors and prevention. Lancet 2006; 367: 1618-25.
[4] Zhao C, Liao Q, Yang D, Yang M and Xu P. Advances in perioperative pain management for total knee arthroplasty: a review of multimodal analgesic approaches. J Orthop Surg Res 2024; 19: 843.
[5] Hannon CP, Fillingham YA, Hamilton WG and Della Valle CJ. Multimodal analgesia and anesthesia: enabling safe and rapid recovery for total joint arthroplasty patients. J Arthroplasty 2022; 37: 1669-1670.
[6] Brummett CM, Waljee JF, Goesling J, Moser S, Lin P, Englesbe MJ, Bohnert ASB, Kheterpal S and Nallamothu BK. New persistent opioid use after minor and major surgical procedures in US adults. JAMA Surg 2017; 152: e170504.
[7] Ward R, Taber D, Gonzales H, Gebregziabher M, Basco W, McCauley J, Mauldin P and Ball S. Risk factors and trajectories of opioid use following total knee replacement. Knee Surg Relat Res 2022; 34: 18.
[8] Johnson A, Milne B, Pasquali M, Jamali N, Mann S, Gilron I, Moore K, Graves E and Parlow J. Long-term opioid use in seniors following hip and knee arthroplasty in Ontario: a historical cohort study. Can J Anaesth 2022; 69: 934-944.
[9] Nair AA, Velagapudi MA, Lang JA, Behara L, Venigandla R, Velagapudi N, Fong CT, Horibe M, Lang JD and Nair BG. Machine learning approach to predict postoperative opioid requirements in ambulatory surgery patients. PLoS One 2020; 15: e0236833.
[10] Klemt C, Harvey MJ, Robinson MG, Esposito JG, Yeo I and Kwon Y. Machine learning algorithms predict extended postoperative opioid use in primary total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc 2022; 30: 2573-2581.
[11] Karhade AV, Ogink PT, Thio QCBS, Cha TD, Gormley WB, Hershman SH, Smith TR, Mao J, Schoenfeld AJ, Bono CM and Schwab JH. Development of machine learning algorithms for prediction of prolonged opioid prescription after surgery for lumbar disc herniation. Spine J 2019; 19: 1764-1771.
[12] Kim SC, Choudhry N, Franklin JM, Bykov K, Eikermann M, Lii J, Fischer MA and Bateman BT. Patterns and predictors of persistent opioid use following hip or knee arthroplasty. Osteoarthritis Cartilage 2017; 25: 1399-1406.
[13] Soffin EM, Wilson LA, Liu J, Poeran J and Memtsoudis SG. Association between sex and perioperative opioid prescribing for total joint arthroplasty: a retrospective population-based study. Br J Anaesth 2021; 126: 1217-1225.
[14] Wilson JM, Yoon J, Mun CJ, Meints SM, Campbell CM, Haythornthwaite JA, Smith MT, Edwards RR and Schreiber KL. The association between changes in clinical pain severity and IL-6 reactivity among patients undergoing total knee arthroplasty: the moderating role of change in insomnia. Brain Behav Immun 2024; 120: 199-207.
[15] Chidambaran V, Duan Q, Pilipenko V, Glynn SM, Sproles A, Martin LJ, Lacagnina MJ, King CD and Ding L. The role of cytokines in acute and chronic postsurgical pain after major musculoskeletal surgeries in a quaternary pediatric center. Brain Behav Immun 2024; 122: 596-603.
[16] Beilin B, Bessler H, Mayburd E, Smirnov G, Dekel A, Yardeni I and Shavit Y. Effects of preemptive analgesia on pain and cytokine production in the postoperative period. Anesthesiology 2003; 98: 151-5.
[17] Nielsen NI, Kehlet H, Gromov K, Troelsen A, Husted H, Varnum C, Kjærsgaard-Andersen P, Rasmussen LE, Pleckaitiene L and Foss NB. High-dose steroids in high pain responders undergoing total knee arthroplasty: a randomised double-blind trial. Br J Anaesth 2022; 128: 150-158.
[18] Shaw JH, Wesemann LD, Banka TR, North WT, Charters MA and Davis JJ. The AAHKS clinical research award: oral dexamethasone following total knee arthroplasty: a double-blind, randomized controlled trial. J Arthroplasty 2023; 38 Suppl 2: S15-S20.
[19] Arraut J, Thomas J, Oakley CT, Barzideh OS, Rozell JC and Schwarzkopf R. The AAHKS best podium presentation research award: a second dose of dexamethasone reduces postoperative opioid consumption and pain in total joint arthroplasty. J Arthroplasty 2023; 38 Suppl 2: S21-S28.
[20] Yuenyongviwat V, Kitjakrancharoensin P and Iamthanaporn CA. Efficacy of weight-based low-dose intravenous dexamethasone for pain management following total knee arthroplasty: a retrospective case-matched study. Orthop Rev (Pavia) 2025; 17: 143092.
[21] Ailioaie RC, Stefanescu E, Leahu C, Boldis A and Scurtu R. The influence of analgesic wound infiltration on postoperative pain and inflammatory cytokines in open colorectal surgery: a randomized comparative pilot study. Medicina (Kaunas) 2024; 60: 1244.
[22] Salimi A, Sabetkasaei M, Raisi H, Labibi F, Ameli H, Khazaei-Poul Y, Zarei M, Mottaghi K, Safari F, Nazem-Bokaei A, Mohajerani S and Moini-Zanjani T. Carbamazepine effects on pain

management and serum IL-6, IL-10 evaluation in addicted patients undergoing surgery. Eur J Pharmacol 2017; 812: 184-188.
[23] Zhuang Q, Tao L, Lin J, Jin J, Qian W, Bian Y, Li Y, Dong Y, Peng H, Li Y, Fan Y, Wang W, Feng B, Gao N, Sun T, Lin J, Zhang M, Yan S, Shen B, Pei F and Weng X. Postoperative intravenous parecoxib sodium followed by oral celecoxib post total knee arthroplasty in osteoarthritis patients (PIPFORCE): a multicentre, double-blind, randomised, placebo-controlled trial. BMJ Open 2020; 10: e030501.
[24] Call CM, Noyes M, Faour KNW, Jeselskis D and Rana AJ. Fewer pills for fewer problems: strategies for reducing postoperative opioid prescribing following total joint arthroplasty. Arthroplast Today 2025; 34: 101758.
[25] Bartley EJ and Fillingim RB. Sex differences in pain: a brief review of clinical and experimental findings. Br J Anaesth 2013; 111: 52-8.
[26] Alexander SN, Green AR, Debner EK, Ramos Freitas LE, Abdelhadi HMK, Szabo-Pardi TA and Burton MD. The influence of sex on neuroimmune communication, pain, and physiology. Biol Sex Differ 2024; 15: 82.
[27] Staehelin Jensen T. The pathogenesis of painful diabetic neuropathy and clinical presentation. Diabetes Res Clin Pract 2023; 206 Suppl 1: 110753.
[28] Hong SH, Kwon SC, Lee JH, Moon S and Kim JI. Influence of diabetes mellitus on postoperative complications after total knee arthroplasty: a systematic review and meta-analysis. Medicina (Kaunas) 2024; 60: 1757.
[29] Xu H, Zhang S, Xie J, Lei Y, Cao G and Pei F. Multiple doses of perioperative dexamethasone further improve clinical outcomes after total knee arthroplasty: a prospective, randomized, controlled study. J Arthroplasty 2018; 33: 3448-3454.
[30] Kurt B, Rex K, Reugels M, Fordyce CB, Fudim M, Sharma A, Berger M, Marx N, Marx-Sch Sh K and Kahles F. Inflammatory biomarkers in heart failure: clinical perspectives on hsCRP, IL-6 and emerging candidates. Curr Heart Fail Rep 2025; 22: 35.
[31] Tanaka T, Narazaki M and Kishimoto T. IL-6 in inflammation, immunity, and disease. Cold Spring Harb Perspect Biol 2014; 6: a016295.
[32] Dominguez E, Rivat C, Pommier B, Mauborgne A and Pohl M. JAK/STAT3 pathway is activated in spinal cord microglia after peripheral nerve injury and contributes to neuropathic pain development in rat. J Neurochem 2008; 107: 50-60.
[33] Dominguez E, Mauborgne A, Mallet J, Desclaux M and Pohl M. SOCS3-mediated blockade of JAK/STAT3 signaling pathway reveals its major contribution to spinal cord neuroinflammation and mechanical allodynia after peripheral nerve injury. J Neurosci 2010; 30: 5754-66.
[34] Kwok CHT, Learoyd AE, Canet-Pons J, Trang T and Fitzgerald M. Spinal interleukin-6 contributes to central sensitisation and persistent pain hypersensitivity in a model of juvenile idiopathic arthritis. Brain Behav Immun 2020; 90: 145-154.
[35] Dai XY, Liu L, Song FH, Gao SJ, Wu JY, Li DY, Zhang LQ, Liu DQ, Zhou YQ and Mei W. Targeting the JAK2/STAT3 signaling pathway for chronic pain. Aging Dis 2024; 15: 186-200.
[36] Gabriel RA, Harjai B, Prasad RS, Simpson S, Chu I, Fisch KM and Said ET. Machine learning approach to predicting persistent opioid use following lower extremity joint arthroplasty. Reg Anesth Pain Med 2022; 47: 313-319.
[37] Yeh PH, Yang SF, Huang JY and Yeh CB. Prediction model for factors associated with long-term opioid use following total knee arthroplasty: a retrospective population-based study. Eur J Pain 2026; 30: e70195.
[38] Vickers AJ and Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006; 26: 565-74.
[39] Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr and Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008; 27: 157-172; discussion 207-212.