



## OPEN ACCESS

## EDITED BY

Alessandra Cuomo,  
University of Naples Federico II, Italy

## REVIEWED BY

Joseph Girard Nugent,  
Oregon Health & Science University,  
United States  
Wanying Su,  
Hunan Provincial People's Hospital, China  
Swapna Gokhale,  
Monash University, Australia

## \*CORRESPONDENCE

Xiao-quan Wang  
✉ tmu\_wxq@126.com  
Jun Liu  
✉ liujun1968tju@163.com

†These authors have contributed equally to  
this work

RECEIVED 05 September 2025

ACCEPTED 21 October 2025

PUBLISHED 05 November 2025

## CITATION

Zhang Y-x, He S, Yang T, Li H-l, Wu C-l,  
Wang L, Wang X-q and Liu J (2025) A  
machine learning approach to predict  
postoperative sleep disturbance after total  
knee arthroplasty: a comparative study  
of multiple algorithms.  
*Front. Med.* 12:1699842.  
doi: 10.3389/fmed.2025.1699842

## COPYRIGHT

© 2025 Zhang, He, Yang, Li, Wu, Wang, Wang  
and Liu. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The  
use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# A machine learning approach to predict postoperative sleep disturbance after total knee arthroplasty: a comparative study of multiple algorithms

Yi-xiang Zhang<sup>1,2†</sup>, Sen He<sup>3†</sup>, Tao Yang<sup>1,2†</sup>, Hao-liang Li<sup>1,2</sup>,  
Chun-lei Wu<sup>1,4</sup>, Lei Wang<sup>2</sup>, Xiao-quan Wang<sup>2\*</sup> and Jun Liu<sup>1,2\*</sup>

<sup>1</sup>College of Orthopedics, Tianjin Medical University, Tianjin, China, <sup>2</sup>Department of Joints, Tianjin Hospital of Tianjin University (Tianjin Hospital), Tianjin, China, <sup>3</sup>Department of Joints, Beichen Hospital, Tianjin, China, <sup>4</sup>The United Logistics Support Force of 983 Hospital, Tianjin, China

**Background:** Postoperative sleep disturbance (PSD) is a common complication following total knee arthroplasty (TKA), which negatively impacts patient recovery. Despite the critical need for early detection and management, there is limited research on predictive models for early PSD, particularly those integrating machine learning (ML) techniques.

**Objective:** This study aimed to develop a predictive model for early PSD following TKA using ML algorithms, identify key predictive factors, and provide an interpretable model to guide clinical decision-making.

**Methods:** The study included 505 patients who underwent TKA. Clinical data were collected at three stages: preoperatively, intraoperatively, and postoperatively. Ten ML models, including logistic regression, support vector machine (SVM), and XGBoost, were trained and evaluated using a test set. Performance metrics, including accuracy, sensitivity, specificity, and area under the curve (AUC), were used to evaluate the efficacy of the models. Key features influencing PSD were identified through SHapley Additive Explanations (SHAP) analysis to enhance model interpretability.

**Results:** Gradient Boosting Machine (GBM) demonstrated the highest AUC (0.906), accuracy (0.834), and sensitivity (0.879), establishing it as the optimal model for predicting PSD. Key predictors identified included age, smoking, living alone, living in the city, VAS 1 month postoperative, and anxiety 1 month postoperative. SHAP analysis revealed that postoperative VAS and age were the most influential factors in predicting PSD, with their impact varying based on individual patient data.

**Conclusion:** The study developed a robust and interpretable ML model for the early prediction of PSD following TKA. This model can aid in preoperative risk stratification, facilitating personalized management strategies to improve postoperative outcomes. Further validation in larger cohorts and diverse settings is necessary to enhance its broader clinical applicability.

## KEYWORDS

postoperative sleep disturbance, total knee arthroplasty, machine learning, predictive model, SHAP analysis

## Introduction

Knee osteoarthritis (OA) is a common musculoskeletal disorder that causes significant pain and dysfunction (1). TKA is an effective treatment for end-stage knee OA, and its use is increasing due to the aging population and advancements in surgical techniques (2–4). TKA is highly effective in relieving pain, improving joint function, and enhancing quality of life; however, up to 20% of patients remain dissatisfied after the procedure (5–8). The etiology of patient dissatisfaction after TKA is multifactorial. Sleep disorders, particularly postoperative sleep disturbance (PSD), are increasingly recognized as common and detrimental to recovery after TKA (9). These disorders can negatively affect postoperative pain management, mental health, and overall recovery (10, 11).

Postoperative sleep disturbance is a common but often underrecognized complication following surgical procedures. Studies suggest that the incidence of perioperative PSD in patients undergoing TKA may exceed 50% (12). PSD is characterized by difficulty falling asleep, fragmented sleep, frequent nocturnal awakenings, and poor sleep quality (9). PSD not only causes patient dissatisfaction but also has significant negative consequences on postoperative recovery (13). PSD is strongly associated with worsened pain perception, complicating postoperative pain management and delaying recovery (14–16). Additionally, PSD is linked to increased levels of anxiety and depression, further hindering rehabilitation and prolonging recovery (17). Furthermore, poor sleep quality impairs immune function, delays wound healing, and increases the risk of postoperative complications, resulting in longer hospital stays and higher healthcare costs (18–20).

Current research on PSD predictors primarily relies on traditional statistical approaches, particularly logistic regression (21). While valuable, these methods depend on pre-specified linear assumptions and struggle to capture the complex, non-linear interactions among numerous clinical, psychological, and social factors influencing postoperative sleep. This limitation necessitates analytical approaches that can automatically learn these complex patterns from data. ML has shown remarkable potential in this regard, revolutionizing predictive modeling across various medical specialties with its robust data processing capabilities and superior predictive performance (22, 23). Despite these advancements, however, the application of ML to PSD prediction following TKA remains underdeveloped, with a notable scarcity of dedicated models in the current literature.

This study aims to address this critical research gap by developing and validating a comprehensive ML-based predictive model for PSD following TKA. Our investigation includes preoperative, intraoperative, and early postoperative clinical data to identify key predictive factors. A fundamental innovation of our approach is the integration of novel socio-environmental predictors, such as "living alone" and "urban residence," which have been largely overlooked in previous research despite their potentially significant impact on sleep quality during recovery. Through interpretability analysis, we identify key predictive factors for PSD occurrence. Our findings are expected to enable early identification of at-risk patients, support preoperative risk stratification, improve perioperative management, and ultimately facilitate personalized rehabilitation strategies after TKA.

## Materials and methods

### Study design and patient selection

This study was approved by the Institutional Review Board (IRB) of Tianjin Hospital (IRB 2024 Medical Ethics Review 102). All procedures involving human participants were conducted in accordance with the ethical standards established by the IRB and the Declaration of Helsinki. All participants signed informed consent forms, explicitly stating that their clinical data would be used for research and model development. Additionally, all data were de-identified during use to ensure patient privacy and security.

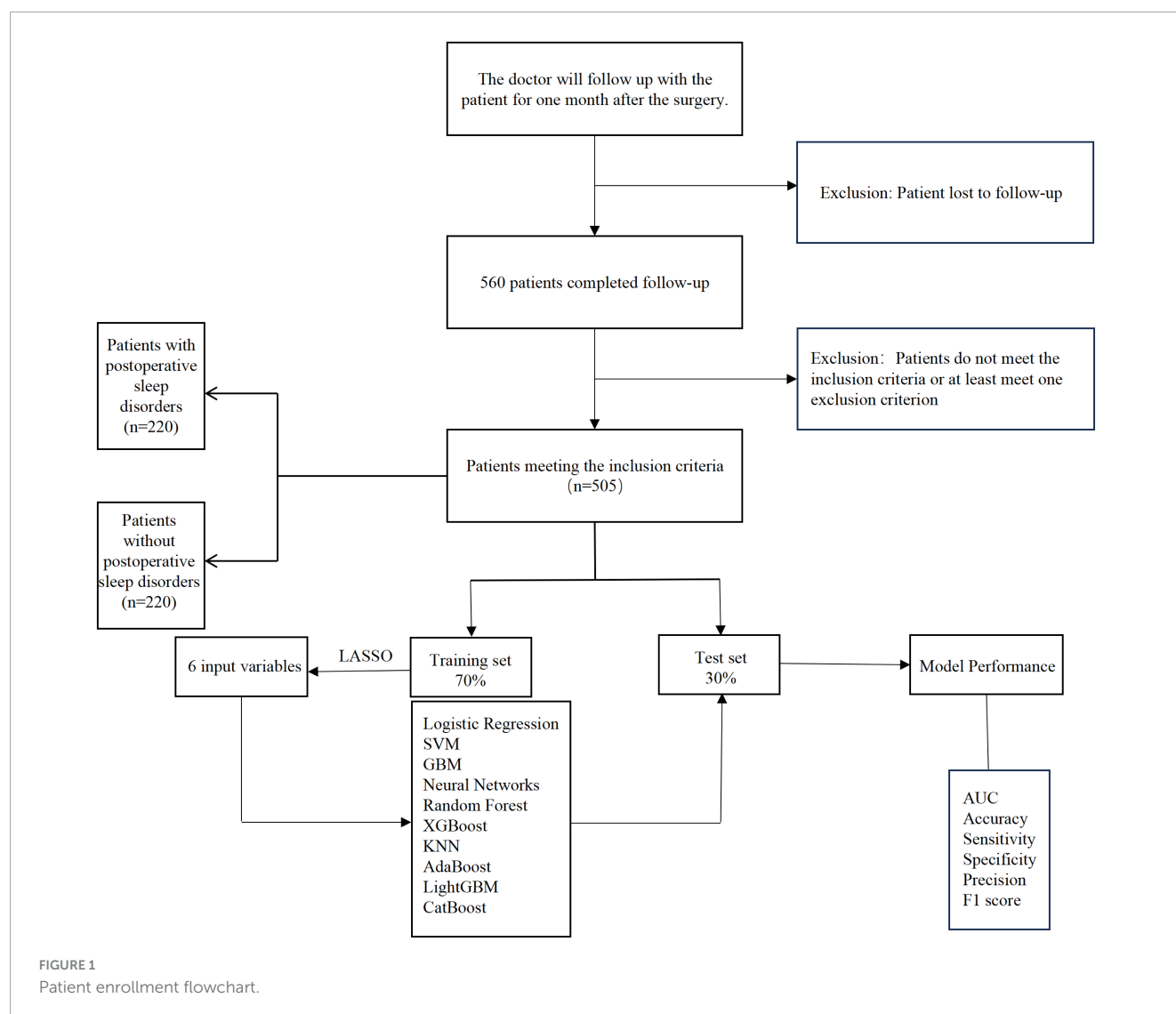
This study included patients who underwent TKA at the Department of Joint Surgery, Tianjin Hospital, between May 2024 and March 2025 for retrospective analysis. The Pittsburgh Sleep Quality Index (PSQI), a widely used self-assessment tool for sleep evaluation, reflects sleep status and quality over the past month. A total PSQI score above 5 indicates poor sleep quality (24). A previous study reported that the incidence of sleep disturbance at 4 weeks postoperatively was 31% (25). The follow-up period in this study was 1 month postoperatively, with the presence of PSD defined by a PSQI score greater than 5.

Importantly, while the  $PSQI > 5$  was used to define the presence of sleep disturbance, this threshold was not used as an inclusion criterion for the study. Instead, all patients who underwent TKA between May 2024 and March 2025 were included in the study regardless of their PSQI scores. Following the application of the exclusion criteria, 505 patients were included in the analysis, with 220 patients diagnosed with PSD ( $PSQI > 5$ ) and 285 patients not diagnosed with PSD ( $PSQI \leq 5$ ) (Figure 1).

The inclusion criteria were: primary osteoarthritis in patients aged 50–80 years undergoing unilateral TKA. The exclusion criteria included: (1) patients with preoperative sleep disturbances ( $PSQI > 5$ ), (2) severe cognitive or psychiatric disorders, (3) regular use of sleep aids during the perioperative period, (4) prior treatment with other systemic psychological interventions, and (5)  $>20\%$  missing clinical data.

### Data collection and data preprocessing

The majority of the data were derived from the electronic patient record (ePR) system at Tianjin Hospital and its associated Clinical Data Analysis and Reporting System (CDARS), with the remaining data obtained from postoperative follow-up. As this was a retrospective study, PSQI scores were collected as part of routine clinical care at preoperative visits and postoperative follow-ups, not prospectively assessed specifically for research purposes. Data with more than 20% missing values were excluded from the analysis (26). A total of 38 variables were analyzed, including demographic data (e.g., age, gender, smoking, alcohol consumption, medical history), laboratory results (e.g., WBC, HB, CR, TP), and 1-month postoperative follow-up data [e.g., visual analogue scale (VAS), WOMAC, anxiety levels]. These variables were selected based on clinical plausibility to form a comprehensive feature set for data-driven prediction modeling of PSD.



Patient-reported outcomes and functional measures, including the VAS for pain (27), the Self-Rating Anxiety Scale (SAS) (28), the Self-Rating Depression Scale (SDS) (29), the WOMAC score (30), and knee range of motion, were assessed and documented by experienced clinicians at four postoperative time points: days 7, 14, 21, and 28 during routine follow-up visits. Assessments were conducted using standardized, validated tools. For analysis, the arithmetic mean of the four measurements was computed for each variable to obtain a representative “1-month postoperative” value. This approach was adopted to improve the reliability of the measurement by reducing the influence of daily fluctuations, thereby offering a more stable estimate of the patient’s typical state during the recovery period. An SAS score >50 indicated mild anxiety, and an SDS score >53 indicated mild depression (28, 29). These instruments were widely recognized and validated in clinical practice. This approach ensured data accuracy and reliability, with evaluations conducted by trained healthcare professionals.

The subsequent data cleaning and preprocessing steps involved standardization and conversion of text descriptions into numerical values to ensure dataset quality and accuracy. Continuous variables were retained in their original form. Binary variables, such as

gender, were coded (female = 0, male = 1). PSD patients were classified as “cases,” while non-PSD patients were classified as “controls,” with respective coding of 1 and 0. Missing data for continuous variables were imputed using the expectation-maximization method. Missing values for binary variables were imputed using the mode (Supplementary Table 1). Only variables with missing data less than 20% were imputed, while large amounts of missing data were excluded during the patient selection phase (26). This approach ensured the model was developed with a complete, reliable dataset, without artificially inflating the sample size. The characteristics of the data were summarized in Table 1.

## Statistical analyses and model development

This study began with data preparation and anonymization, followed by preliminary cleaning, which involved removing duplicates and imputing missing values. To develop the predictive model, all preprocessed variables were incorporated directly into

TABLE 1 Characteristics of the cohort.

| Variables                                    | Non-PSD        | PSD            | P-values     |
|----------------------------------------------|----------------|----------------|--------------|
| <b>Demographics</b>                          |                |                |              |
| Age (years)                                  | 71.2 ± 8.0     | 72.4 ± 4.9     | <b>0.023</b> |
| Sex, <i>n</i> (%)                            |                |                | 0.821        |
| Female                                       | 196            | 151            |              |
| Male                                         | 88             | 70             |              |
| Smoking, <i>n</i> (%)                        | 88 (30.8%)     | 79 (35.9%)     | <b>0.037</b> |
| Drinking, <i>n</i> (%)                       | 65 (22.8%)     | 51 (23.2%)     | 0.921        |
| Diabetes, <i>n</i> (%)                       | 116 (40.7%)    | 94 (42.7%)     | 0.430        |
| Hyperlipidemia, <i>n</i> (%)                 | 109 (38.2%)    | 84 (38.1%)     | 0.988        |
| Hypertension, <i>n</i> (%)                   | 110 (38.6%)    | 87 (39.5%)     | 0.828        |
| Complication, <i>n</i> (%)                   | 96 (33.7%)     | 70 (31.8%)     | 0.658        |
| Live alone                                   | 28(9.8%)       | 53(24%)        | <b>0.013</b> |
| Live in the city                             | 126(44.2%)     | 118(53.6%)     | <b>0.022</b> |
| <b>Laboratory findings</b>                   |                |                |              |
| WBC (10 <sup>9</sup> /L)                     | 5.95 ± 1.16    | 6.10 ± 0.29    | 0.055        |
| HB (g/L)                                     | 132.89 ± 11.99 | 133.11 ± 10.47 | 0.780        |
| Cr (mg/dL)                                   | 79.65 ± 15.72  | 80.20 ± 18.16  | 0.622        |
| TP (g/L)                                     | 69.91 ± 5.60   | 67.30 ± 3.00   | 0.347        |
| ALB (g/L)                                    | 36.63 ± 3.11   | 40.10 ± 3.24   | 0.940        |
| UA (μmol/L)                                  | 284.07 ± 80.72 | 285.60 ± 61.09 | 0.648        |
| D_dimer (mg/L)                               | 0.27 ± 0.10    | 0.28 ± 0.11    | 0.302        |
| CRP (mg/L)                                   | 5.31 ± 2.90    | 5.35 ± 0.43    | 0.827        |
| HDL (mmol/L)                                 | 1.33 ± 0.17    | 1.78 ± 0.22    | 0.651        |
| LDL (mmol/L)                                 | 2.82 ± 0.56    | 2.77 ± 0.15    | 0.167        |
| TC (mmol/L)                                  | 4.39 ± 0.38    | 4.40 ± 0.50    | 0.656        |
| ALT (U/L)                                    | 32.55 ± 12.10  | 32.70 ± 11.66  | 0.883        |
| GLB (g/L)                                    | 24.97 ± 1.91   | 24.93 ± 2.10   | 0.821        |
| TB (mmol/L)                                  | 7.50 ± 2.13    | 7.53 ± 2.05    | 0.758        |
| APTT (s)                                     | 30.08 ± 2.33   | 29.91 ± 2.49   | 0.434        |
| PT (s)                                       | 12.45 ± 0.57   | 11.28 ± 0.62   | 0.974        |
| <b>Clinical data</b>                         |                |                |              |
| Preoperative PSQI                            | 4.44 ± 1.25    | 3.69 ± 1.31    | <b>0.035</b> |
| Preoperative anxiety                         | 42.86 ± 7.94   | 42.18 ± 6.12   | 0.987        |
| Preoperative depression                      | 48.10 ± 10.77  | 49.25 ± 5.64   | 0.149        |
| Preoperative VAS                             | 7.07 ± 0.85    | 6.91 ± 1.17    | 0.076        |
| Preoperative Womac                           | 173.79 ± 14.33 | 173.98 ± 8.59  | 0.872        |
| VAS 1 month postoperative                    | 2.30 ± 0.54    | 2.46 ± 1.10    | <b>0.034</b> |
| Womac 1 month postoperative                  | 46.03 ± 4.03   | 45.89 ± 2.92   | 0.558        |
| Anxiety 1 month postoperative                | 37.27 ± 4.78   | 38.32 ± 5.86   | <b>0.027</b> |
| Depression 1 month postoperative             | 15.80 ± 2.99   | 15.94 ± 3.09   | 0.596        |
| One month postoperative knee range of motion | 101.05 ± 8.83  | 105.37 ± 6.18  | <b>0.041</b> |

Bold values indicate statistically significant differences with a *p*-value <0.05. Group comparisons were made using Student's *t*-test for normally distributed continuous variables, Mann-Whitney U test for non-normally distributed continuous variables, and Chi-square test for categorical variables, as detailed in the Section "Statistical analyses and model development".

a Least Absolute Shrinkage and Selection Operator (LASSO) model for training.

Least Absolute Shrinkage and Selection Operator was chosen because it simultaneously performed feature selection and model fitting. The model applied L1 regularization, shrinking the coefficients of less important features to zero, thereby automatically identifying the most influential variables and preventing overfitting (31). This approach avoided biases associated with pre-selection filtering methods and allowed the model to capture complex multivariate relationships.

Candidate variables were initially screened using univariate analysis (*p* < 0.05). The optimal regularization parameter ( $\lambda$ ) was then determined through 10-fold cross-validation, applying the "one standard error" rule (lambda.1se). This criterion selected the most parsimonious model, where the performance was within one standard error of the minimum binomial deviance, thereby favoring model simplicity and robustness.

The dataset was randomly divided into a training set (70%) and a test set (30%) based on common practices in predictive modeling. While this approach was widely used, alternative techniques such as bootstrapping or cross-validation could be considered in future studies to further validate the robustness of the model. The training set was used for model development and hyperparameter optimization, whereas the independent test set was reserved solely for the final evaluation of model performance. For model development, we employed ten ML algorithms: Logistic Regression, support vector machine (SVM), Gradient Boosting Machine (GBM), Neural Networks, Random Forest, eXtreme Gradient Boosting (XGBoost), K-Nearest Neighbors (KNN), AdaBoost, Light Gradient Boosting Machine (LightGBM), and Categorical Boosting (CatBoost). These ten models were selected to ensure comprehensive coverage of major, high-performing machine learning families, including linear models, support vector machines, tree-based ensembles, and boosting algorithms (32). This approach guaranteed a robust and representative comparison of state-of-the-art techniques applicable to structured clinical data. While deep learning approaches were considered, they were not adopted due to the moderately-sized dataset, which was suboptimal for training complex deep networks, and our emphasis on model interpretability for potential clinical use.

Each model was trained using 10-fold cross-validation to assess performance, and hyperparameters were optimized using Bayesian optimization to improve predictive accuracy. The performance of all models was evaluated at each iteration using multiple metrics: AUC, accuracy, sensitivity, specificity, and F1 scores. AUC was prioritized as the primary evaluation metric because it provides a more comprehensive measure of model discrimination, especially in imbalanced datasets (33). AUC represents the area under the curve plotting the true positive rate against the false positive rate, reflecting the model's predictive ability. The AUC ranges from 0 to 1. Models with an AUC greater than 0.7 are considered to exhibit good performance and clinical significance, with an AUC of 1 representing perfect performance (34). For the remaining metrics, values range from 0 to 1, with higher scores indicating better performance. Given the imbalanced nature of the classification task, AUC and balanced accuracy were emphasized during performance evaluation. The average score across iterations determined each model's final performance. Among the ten models, the one with the highest AUC was selected as the final model.

To enhance the transparency and interpretability of the final predictive model, both global and local interpretations were incorporated. The global interpretation was presented using the SHapley Additive Explanations (SHAP) summary plot, while local interpretations were visualized with SHAP waterfall plots for individual PSD cases following TKA (35, 36). According to the SHAP legend, the larger the absolute value of a SHAP value in the waterfall plot, the greater its impact on the prediction. Furthermore, differences in performance for the same feature across individuals, as shown in the single-sample waterfall plots, may have arisen from individual variability, highlighting the model's ability to capture subject-specific differences. We then compared the comprehensive performance metrics of the GBM model across key patient subgroups, with particular focus on socio-environmental predictors and gender distribution. Through these subgroup analyses, we aimed to specifically assess potential model bias and better understand the model's applicability across different patient demographics, thereby providing evidence for its fairness and generalizability. Since different subgroups may have experienced varying degrees of class imbalance which can significantly impact model performance (37), we evaluated multiple metrics including Accuracy, Sensitivity, Specificity, Precision, F1-score, and the AUC to thoroughly assess the model's performance in these specific populations. The comprehensive analysis provided valuable insights into model fairness and offered targeted data support for personalized treatment strategies.

This study described the characteristics of various datasets and conducted a series of statistical tests. For continuous data, means and standard deviations were used for normally distributed variables, while medians and interquartile ranges were applied to non-normally distributed variables. Categorical data were summarized using frequencies and proportions. Group comparisons were made using the Student's *t*-test for normally distributed continuous variables, the Mann-Whitney U test for non-normally distributed continuous variables, and the Chi-square test for categorical variables. A two-tailed *p*-value of  $< 0.05$  was considered statistically significant. All statistical analyses and model construction were performed using IBM SPSS Statistics (version 26.0) and R (version 4.4.2).

## Results

### Cohort characteristics

This study included 505 patients, of whom 220 were diagnosed with PSD and 285 were classified as normal. The prevalence of PSD in our cohort was 43.6%. This finding aligns with the established literature, highlighting the substantial burden of this complication in the postoperative period (12). Among the total patient population, 347 (68.7%) were female, and 158 (31.3%) were male. The mean age of the patients was  $71.7 \pm 6.9$  years, with a mean BMI of  $22.4 \pm 4.4$ . A total of 167 patients had a history of smoking, and 116 patients had a history of alcohol consumption. Among comorbidities, diabetes mellitus was the most common, affecting 210 patients (41.6%), followed by hypertension in 197 patients (39.0%) and hyperlipidemia in 193 patients (38.2%). Regarding patient residence, 81 patients (16.0%) lived alone, and

244 patients (48.3%) lived in the city. The baseline demographics, along with the results of univariate and multivariate analyses, are presented in Table 1.

### Predictors screened by LASSO regression

Using PSD as the dependent variable, LASSO regression with 10-fold cross-validation identified six key predictors from the initial candidate variables: smoking, age, VAS 1 month postoperative, anxiety 1 month postoperative, living alone, and living in the city (Figures 2A, B). These findings highlight the key risk factors associated with the development of PSD in post-TKA patients, which can assist in clinical decision-making and guide targeted interventions.

### Model performance

The performance of ten ML models was evaluated on the test set, with AUC values ranging from 0.666 to 0.906. Among these models, the Logistic model demonstrated the lowest AUC, while the GBM model achieved the highest AUC, indicating superior discriminative ability. In terms of accuracy, the Logistic model had the lowest value at 0.675, while the XGBoost model achieved the highest accuracy at 0.874. For sensitivity, the AdaBoost model scored the lowest at 0.576. The GBM and Random Forest models achieved the highest sensitivity score of 0.879. For specificity, the LightGBM model performed best, achieving a specificity of 0.906. For precision, the LightGBM model achieved the highest score of 0.864. For the F1 score, the Logistic model scored the lowest at 0.647, while the Random Forest model achieved the highest score at 0.853. Overall, the GBM model demonstrated the best discriminative ability among all ten models and performed consistently and reliably during 10-fold cross-validation. Therefore, the GBM model was selected as the final prediction model (Figure 3 and Table 2).

### Feature importance

SHapley Additive Explanations summary plots offered a global interpretation of model decisions, visualizing the importance of each feature (Figures 4, 5). This analysis confirmed the importance of the six LASSO-selected predictors and further quantified their effects. The model identified VAS 1 month postoperative and age as the most influential factors, followed by anxiety 1 month postoperative, living alone, urban residence, and smoking. Overall, all identified predictors were risk factors for PSD post-TKA.

We provided two localized SHAP waterfall plots for individual patients to illustrate patient-level interpretations of the final model predictions (Figures 6, 7). Figure 6 shows the 28th TKA patient in our cohort. In this case, VAS 1 month postoperative (2.6) was the most significant risk factor, followed by anxiety 1 month postoperative (38) and living alone. Not smoking and not living in the city were the most important protective factors. Figure 7 shows the 35th TKA patient in our cohort. In this case, VAS 1 month postoperative (2.9) was the most significant risk factor, followed by

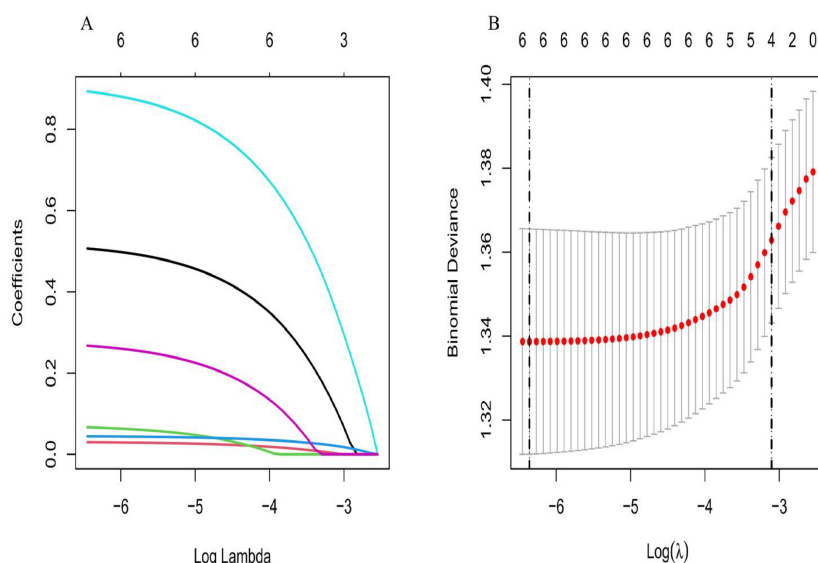


FIGURE 2

(A) Least Absolute Shrinkage and Selection Operator (LASSO) coefficient path plot: This plot shows how the coefficients of different features change as the  $\lambda$  value increases in a LASSO regression model. As  $\lambda$  increases, the coefficients of less important features are progressively compressed toward zero. Features that reach zero early contribute less to the model, while those that remain non-zero for a longer period are more influential, indicating their greater relevance in the prediction task. The plot helps in visualizing which features are selected and retained in the model as regularization strength increases. (B) Cross-validation curve for LASSO regression: The plot illustrates the binomial deviance (model error) as a function of  $\log(\lambda)$  in a LASSO regression model. The solid curve represents the mean binomial deviance, and the shaded area between the dashed lines indicates the range of one standard deviation above and below the mean. The optimal value of  $\log(\lambda)$  is determined where the error is minimized, corresponding to the lowest deviance, as indicated by the vertical dashed lines. This curve aids in selecting the best regularization parameter for minimizing model error.

anxiety 1 month postoperative (39), living alone, and living in the city. Not smoking was the most important protective factor.

## Subgroup analysis

We conducted a detailed subgroup analysis of the final GBM model to evaluate its fairness and generalizability, focusing on social environment and gender factors (Table 3). The model demonstrated strong and consistent predictive performance across most subgroups, with AUC values consistently above 0.88 in gender-based (male/female) and urban residence subgroups. However, performance showed variability in the “living alone” subgroup ( $N = 81$ ). This fluctuation is likely due to the small sample size in this subgroup, which limited the model’s ability to identify stable patterns, combined with a disproportionately high percentage of PSD patients (65.4%), which exacerbated the impact of class imbalance on model stability. These findings suggest that while the model performs reliably overall, caution is warranted when applying it to patients living alone. Future validation with larger sample sizes is necessary to confirm these results.

## Discussion

This study aims to develop an ML-based model for predicting PSD in patients following TKA. A key innovation of our study is the integration of ten different machine learning models, which offer a multidimensional and comprehensive analytical framework

to predict and identify the main risk factors for postoperative sleep disorders. Using machine learning models, we identify two socio-environmental factors—living alone and living in the city—as predictors for the first time, factors that have not received adequate attention in the existing literature. In addition to these two newly identified factors, our study further confirms the importance of clinical factors, such as VAS scores, anxiety symptoms, and age, in predicting PSD.

## Discovery of innovative socio-environmental factors

This study is the first to highlight the significant role of two factors—living alone and living in the city—in predicting PSD. Patients living alone lack care and assistance from family members after surgery, presenting additional challenges during their recovery. The absence of family support, particularly during the postoperative recovery period, often makes it difficult for these patients to manage pain, perform daily activities, and access necessary psychological support (38, 40, 41). Patients living alone are more likely to feel isolated and anxious, and this emotional burden may exacerbate their sleep disorders (39, 42–45). Therefore, living alone is not only a sociological factor but also reflects the vulnerability of patients’ quality of life and postoperative recovery.

Patients living in urban areas are exposed to a range of environmental stressors, including noise pollution, light pollution, and air pollution (46–48). These environmental factors can affect patients’ sleep quality in several ways, particularly during

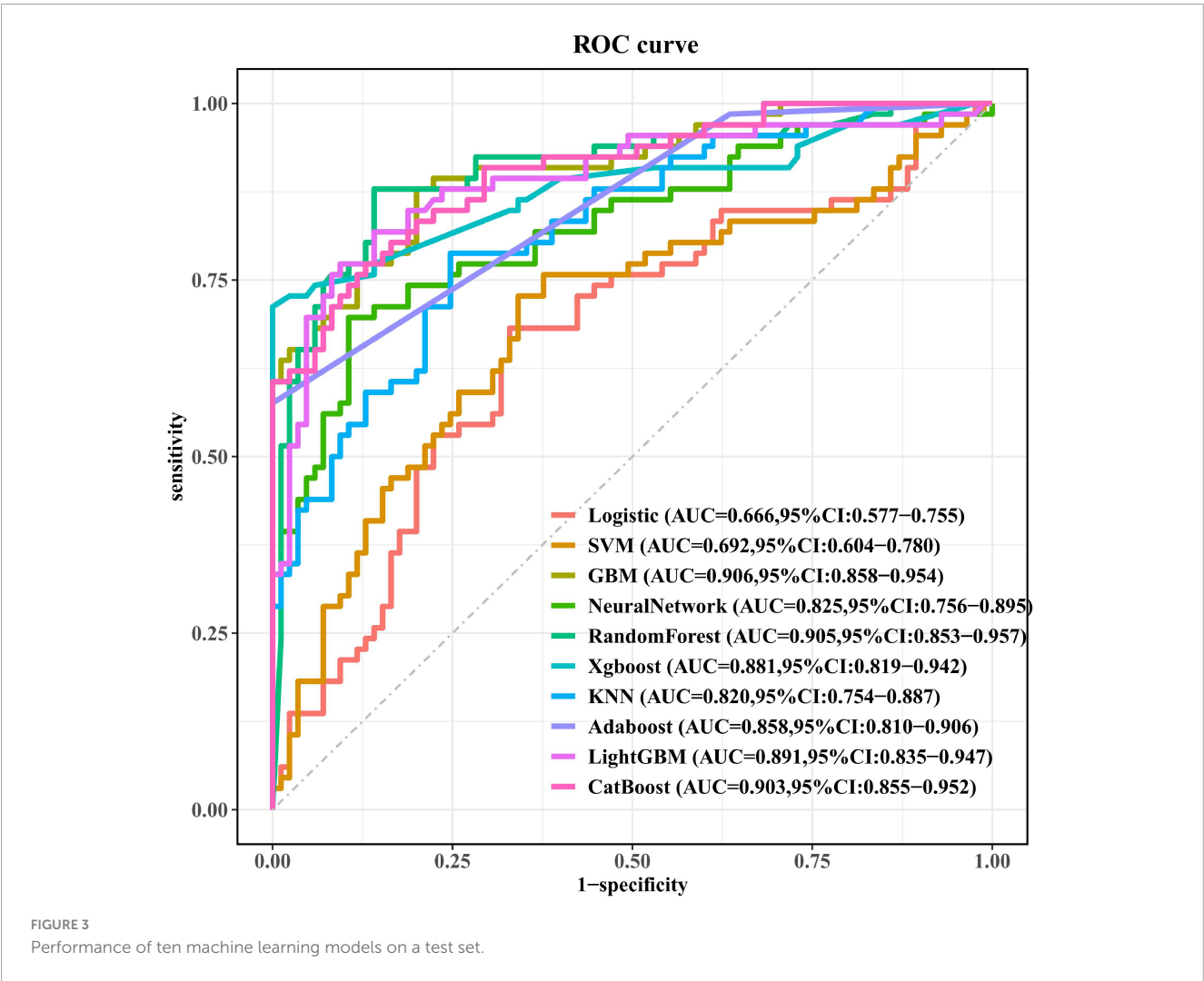


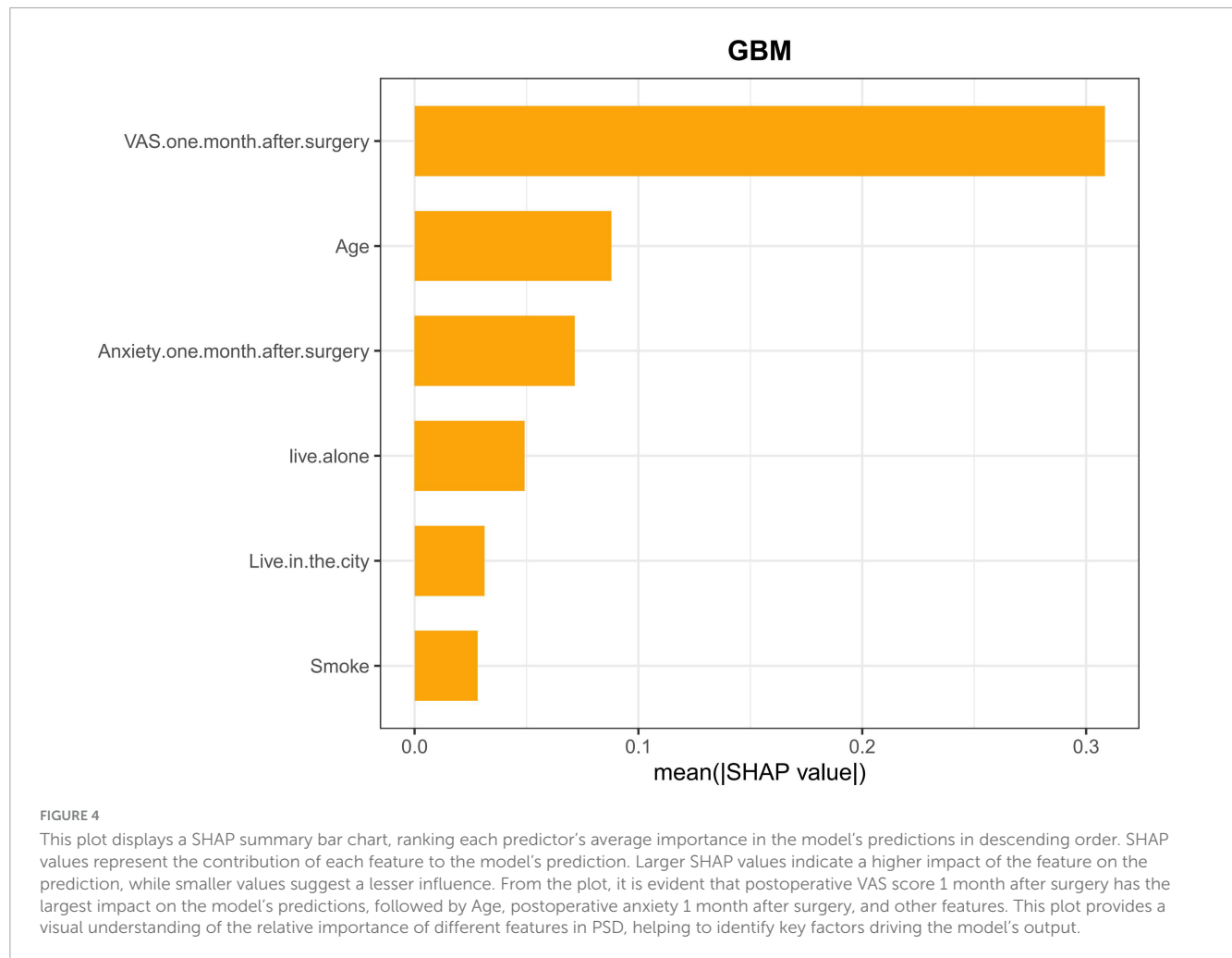
TABLE 2 Performance metrics of different machine learning models for predicting PSD.

| Model          | Accuracy | Sensitivity | Specificity | Precision | F1    |
|----------------|----------|-------------|-------------|-----------|-------|
| Logistic       | 0.675    | 0.682       | 0.671       | 0.616     | 0.647 |
| SVM            | 0.689    | 0.727       | 0.659       | 0.623     | 0.671 |
| GBM            | 0.834    | 0.879       | 0.8         | 0.773     | 0.823 |
| Neural Network | 0.808    | 0.697       | 0.894       | 0.836     | 0.76  |
| Random Forest  | 0.868    | 0.879       | 0.859       | 0.829     | 0.853 |
| Xgboost        | 0.874    | 0.712       | 0.871       | 0.652     | 0.832 |
| KNN            | 0.768    | 0.788       | 0.753       | 0.712     | 0.748 |
| Adaboost       | 0.815    | 0.576       | 0.775       | 0.717     | 0.731 |
| LightGBM       | 0.848    | 0.773       | 0.906       | 0.864     | 0.816 |
| CatBoost       | 0.828    | 0.773       | 0.871       | 0.823     | 0.797 |

the postoperative recovery phase (49, 50). Higher noise levels and light pollution in urban areas may decrease sleep quality, disrupt biological clocks and sleep cycles, and increase the risk of PSD (51–54). Additionally, air pollution and the urban heat island effect may slow the recovery process and increase the incidence of postoperative complications (55, 56). Therefore, the

living environment plays a significant moderating role in the development of PSD after TKA.

The findings of these social and environmental factors highlight that social support and environmental conditions are just as important as medical treatment during postoperative recovery. Therefore, these factors should be considered when



developing postoperative interventions to ensure a more personalized care strategy.

## Validation of clinical factors and the benefits of machine learning models

Besides the two innovative factors—living alone and living in the city—our study also confirmed the role of traditional clinical factors in predicting postoperative sleep disorders. For example, the VAS score (postoperative pain score) is a significant risk factor for PSD. High VAS scores are associated with poorly managed postoperative pain, and persistent pain not only affects sleep quality but may also impact mood and recovery (57–60). Therefore, managing postoperative pain is crucial to reducing the risk of PSD.

Additionally, postoperative anxiety scores are identified as significant predictors. Postoperative anxiety exacerbates patients' pain perception and affects their psychological state, thereby increasing the incidence of sleep disorders. Our study finds that anxiety symptoms are strongly associated with sleep disorders, indicating the need for effective management of anxiety symptoms in postoperative patients to reduce the risk of sleep disorders.

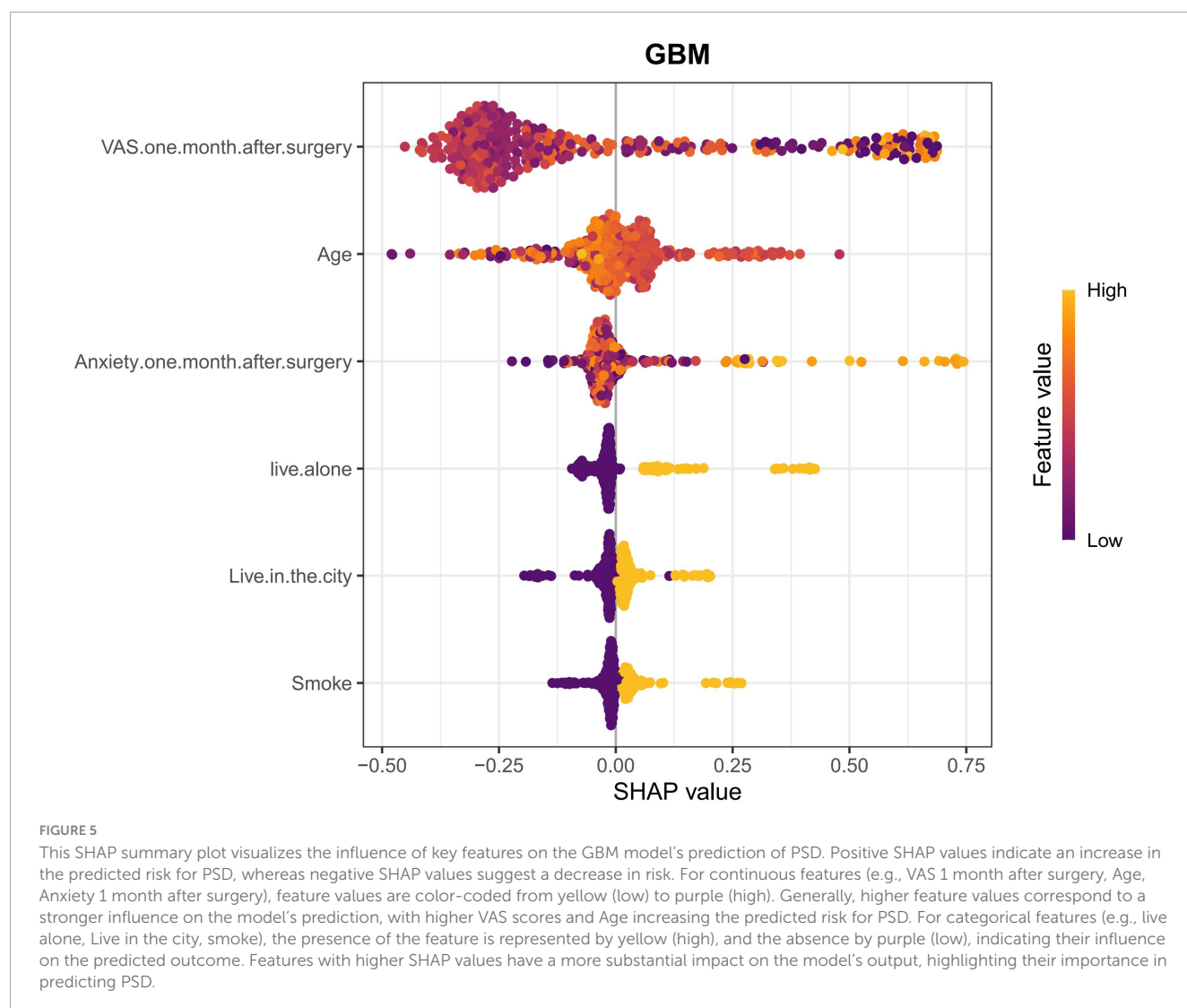
Age is another known influencing factor, as patients' physiological conditions and rehabilitation capacity change

with age. Older patients are at higher risk for comorbidities, such as hypertension and diabetes, which increase the incidence of postoperative sleep disorders (61–63). Our findings confirm the importance of age in postoperative sleep disorders and suggest that elderly patients require special attention for postoperative care and sleep health.

## Applications and benefits of machine learning models

Another innovation in this study is the use of ten machine learning models to analyze the data, including Logistic Regression, SVM, GBM, and Random Forest. Compared to traditional statistical methods, ML handles non-linear relationships and extracts key factors from complex multidimensional data. Through the comparative evaluation of these models, we identify the GBM as the best performer, with high accuracy and sensitivity.

Our research highlights the significant potential of ML in medical prediction, particularly for complex health issues like PSD. By integrating various ML algorithms, we can accurately identify high-risk patients for PSD and offer personalized clinical intervention recommendations. For instance, using our model, clinicians can identify high-risk patients early and implement



appropriate management strategies, such as pain control, anxiety management, and adjustments to the living environment.

## Clinical application and deployment

The findings of this study have significant implications for clinical practice. The predictive model can be integrated into the preoperative assessment process for TKA patients. By using readily available clinical and social data, clinicians can identify patients at high risk for PSD prior to surgery, enabling proactive and personalized management strategies. For example, high-risk patients can be referred to prehabilitation programs focused on pain and anxiety management and offered counseling on sleep hygiene. Postoperatively, these patients can be monitored more closely, and non-pharmacological interventions (e.g., minimizing nighttime disruptions, cognitive behavioral therapy for insomnia) can be started early.

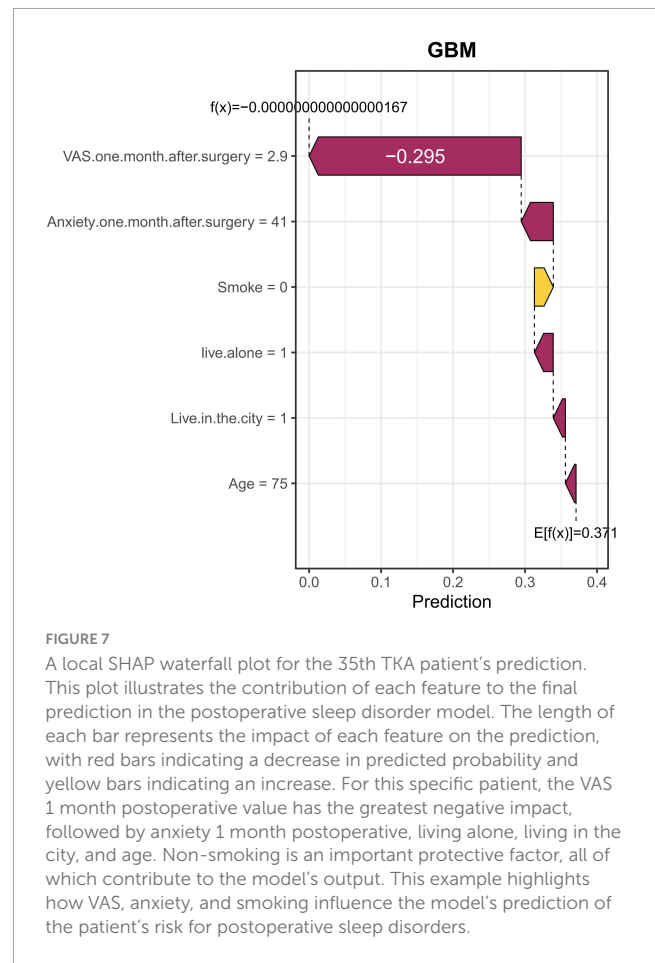
Importantly, our model identifies modifiable risk factors, such as postoperative VAS and anxiety, suggesting that PSD is a largely preventable complication. The model should not be viewed as a deterministic prognosis but as a tool for risk stratification that

identifies specific areas for intervention. By effectively managing pain and addressing anxiety during the perioperative period, the incidence and severity of PSD can be significantly reduced. This model represents a shift from reactive treatment to proactive prevention, providing a pathway for improving postoperative care and outcomes.

Although the ML model shows promise, its successful deployment in clinical practice requires several key considerations. First, the model must be integrated into existing clinical workflows and decision-making systems to facilitate its use by healthcare professionals. Training and adaptation to various clinical settings are essential for effective use.

From a technical standpoint, the model should be scalable and capable of processing large volumes of patient data in real-time without excessive computational requirements. It is also crucial to validate the model across various hospitals and patient populations to ensure its generalizability and applicability.

While this study offers significant innovative value, several limitations should be considered to contextualize the findings and guide future research. First, the single-center, retrospective design, while providing a robust initial dataset, may limit the



1 month limits our understanding of the model's ability to predict persistent sleep disturbances. Future studies incorporating these omitted factors and longer-term outcomes are crucial for enhancing the model's comprehensiveness and clinical relevance. Finally, in terms of model evaluation, our analysis primarily focuses on discriminative performance (the ability to distinguish between PSD and non-PSD patients). We do not formally assess model calibration, which measures the accuracy of predicted risk probabilities. As calibration is a key metric for evaluating the clinical usefulness of a predictive model, investigating it remains

TABLE 3 Performance of the final GBM model across different patient subgroups.

| Subgroup         | AUC   | Accuracy | Sensitivity | Specificity | Precision | F1    |
|------------------|-------|----------|-------------|-------------|-----------|-------|
| Live in the city |       |          |             |             |           |       |
| Yes              | 0.910 | 0.840    | 0.885       | 0.805       | 0.780     | 0.829 |
| No               | 0.901 | 0.828    | 0.873       | 0.802       | 0.766     | 0.816 |
| Live alone       |       |          |             |             |           |       |
| Yes              | 0.870 | 0.790    | 0.820       | 0.770       | 0.740     | 0.778 |
| No               | 0.912 | 0.842    | 0.885       | 0.810       | 0.782     | 0.883 |
| Gender           |       |          |             |             |           |       |
| Female           | 0.902 | 0.831    | 0.875       | 0.798       | 0.770     | 0.819 |
| Male             | 0.911 | 0.839    | 0.884       | 0.805       | 0.781     | 0.829 |

an important area for future work. Additionally, despite our efforts to conduct subgroup analyses, potential biases arising from imbalances in sociodemographic factors may persist, affecting model performance. Further validation in larger and more diverse populations is recommended to ensure fairness and generalizability.

## Conclusion

This study developed a ML-based model for predicting PSD in patients following TKA. By analyzing factors such as age, smoking history, VAS score, and anxiety score, we identified key predictors of PSD. The GBM model showed the best predictive efficacy, with high accuracy and sensitivity. We further enhanced the model's interpretability using SHAP methodology, enabling clinicians to visualize the specific contribution of each factor to the prediction, facilitating preoperative risk stratification and personalized interventions.

Additionally, our study identified two socio-environmental factors—living alone and living in the city—that have not been sufficiently explored in the literature. Patients living alone face greater postoperative challenges due to lack of family support, while those living in urban areas are more exposed to environmental stressors, such as noise and light pollution, which exacerbate the risk of PSD. These findings offer new insights for clinical interventions, emphasizing the importance of social and environmental factors in postoperative care.

Future studies should validate this model across diverse populations, expand its applicability, and incorporate additional factors such as genetic background and long-term follow-up data to enhance its predictive ability and clinical value.

## Data availability statement

The original contributions presented in this study are included in this article/[Supplementary material](#), further inquiries can be directed to the corresponding authors.

## Ethics statement

The studies involving humans were approved by the Institutional Review Board (IRB) of Tianjin Hospital (IRB 2024 Medical Ethics Review 102). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

## Author contributions

Y-xZ: Data curation, Methodology, Conceptualization, Writing – original draft. SH: Methodology, Conceptualization, Writing – original draft. TY: Writing – original draft, Data curation, Conceptualization. H-IL: Writing – original draft, Project administration, Conceptualization. C-IW: Conceptualization, Writing – original draft, Project administration. LW: Conceptualization, Project administration, Writing – original draft. X-qW: Funding acquisition, Writing – review & editing. JL: Funding acquisition, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research and/or publication of this article. This study was supported by the Tianjin Metrology Science and Technology Project (grant no. 2025TJMT026).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmed.2025.1699842/full#supplementary-material>

## References

- Glyn-Jones S, Palmer A, Agricola R, Price A, Vincent T, Weinans H, et al. Osteoarthritis. *Lancet*. (2015) 386:376–87. doi: 10.1016/S0140-6736(14)60802-3
- Carr A, Robertsson O, Graves S, Price A, Arden N, Judge A, et al. Knee replacement. *Lancet*. (2012) 379:1331–40. doi: 10.1016/S0140-6736(11)60752-6
- Siddiqi A, Levine B, Springer B. Highlights of the 2021 American joint replacement registry annual report. *Arthroplasty Today*. (2022) 13:205–7. doi: 10.1016/j.artd.2022.01.020
- Huang L, Wang Q, Zhao X. A commentary on “incidence, patterns and risk factors for readmission following knee arthroplasty in china: a national retrospective cohort study. *Int J Surg*. (2022) 106:106875. doi: 10.1016/j.ijsu.2022.106875
- Clement N, Burnett R. Patient satisfaction after total knee arthroplasty is affected by their general physical well-being. *Knee Surgery Sports Traumatol. Arthroscopy* (2013) 21:1795. doi: 10.1007/s00167-013-2523-y
- Nam D, Nunley R, Barrack R. Patient dissatisfaction following total knee replacement. *Bone Joint J*. (2014) 96-B:96–100. doi: 10.1302/0301-620X.96B11.34152
- Von Keudell A, Sodha S, Collins J, Minas T, Fitz W, Gomoll A. Patient satisfaction after primary total and unicompartmental knee arthroplasty: an age-dependent analysis. *Knee*. (2014) 21:180–4. doi: 10.1016/j.knee.2013.08.004
- Bourne R, Chesworth B, Davis A, Mahomed N, Charron K. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? *Clin Orthopaedics Related Res*. (2010) 468:57–63. doi: 10.1007/s11999-009-1119-9
- Luo M, Song B, Zhu J. Sleep disturbances after general anesthesia. *Curr Perspect*. (2020) 11:629. doi: 10.3389/fneur.2020.00629
- Paulose J, Wang C, O'Hara B, Cassone V. The effects of aging on sleep parameters in a healthy, melatonin-competent mouse model. *Nat Sci Sleep*. (2019) 11:113–21. doi: 10.2147/nss.S214423
- Terzaghi M, Sartori I, Rustioni V, Manni R. Sleep disorders and acute nocturnal delirium in the elderly: a comorbidity not to be overlooked. *Eur J Internal Med*. (2014) 25:350–5. doi: 10.1016/j.ejim.2014.02.008
- Wylde V, Rooker J, Halliday L, Blom A. Acute postoperative pain at rest after hip and knee arthroplasty: severity, sensory qualities and impact on sleep. *Orthopaedics Traumatol Surg Res*. (2011) 97:139–44. doi: 10.1016/j.otsr.2010.12.003
- Wang X, Hua D, Tang X, Li S, Sun R, Xie Z, et al. The role of perioperative sleep disturbance in postoperative neurocognitive disorders. *Nat Sci Sleep*. (2021) 13:1395–410. doi: 10.2147/nss.S320745
- Liu Y, Xiao S, Yang H, Lv X, Hou A, Ma Y, et al. Postoperative pain-related outcomes and perioperative pain management in china: a population-based study. *Lancet Regional Health Western Pacific*. (2023) 39:100822. doi: 10.1016/j.lanwpc.2023.100822
- Sipilä R, Kalso E. Sleep well and recover faster with less pain—a narrative review on sleep in the perioperative period. *J Clin Med*. (2021) 10:2000. doi: 10.3390/jcm10092000
- O'Gara B, Gao L, Marcantonio E, Subramaniam B. Sleep, pain, and cognition: modifiable targets for optimal perioperative brain health. *Anesthesiology*. (2021) 135:1132–52. doi: 10.1097/aln.0000000000004046
- Chen W, Yang Y, Yang H, Yang D, Qu Y, Yang L, et al. Associations of preoperative sleep disturbance with intraoperative and postoperative adverse outcomes among chinese surgical patients: evidence from the China surgery and anesthesia cohort (Csac). *J Clin Anesthesia*. (2025) 106:111956. doi: 10.1016/j.jclinane.2025.111956
- Knutson K, Spiegel K, Penev P, Van Cauter E. The metabolic consequences of sleep deprivation. *Sleep Med Rev*. (2007) 11:163–78. doi: 10.1016/j.smrv.2007.01.002
- Papathanassoglou E. Psychological support and outcomes for ICU patients. *Nurs Crit Care*. (2010) 15:118–28. doi: 10.1111/j.1478-5153.2009.00383.x
- Cooper A, Hanly P. Sleep and recovery from critical illness and injury: a review of theory, current practice, and future directions. *Crit Care Med*. (2008) 36:2962–3. doi: 10.1097/CCM.0b013e318187268d
- Du J, Zhang H, Ding Z, Wu X, Chen H, Ma W, et al. Development and validation of a nomogram for postoperative sleep disturbance in adults: a prospective survey of 640 patients undergoing spinal surgery. *BMC Anesthesiol*. (2023) 23:154. doi: 10.1186/s12871-023-02097-x
- Bini S. Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care? *J Arthroplasty*. (2018) 33:2358–61. doi: 10.1016/j.arth.2018.02.067
- Lau L, Chui E, Man G, Xin Y, Ho K, Mak K, et al. A novel image-based machine learning model with superior accuracy and predictability for knee arthroplasty loosening detection and clinical decision making. *J Orthopaedic Transl*. (2022) 36:177–83. doi: 10.1016/j.jot.2022.07.004
- Buyse D, Reynolds C, Monk T, Berman S, Kupfer D. The pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res*. (1989) 28:193–213. doi: 10.1016/0165-1781(89)90047-4
- Wang Y, Liu Y, Li X, Lv Q, Xia Q, Wang X, et al. Prospective assessment and risk factors of sleep disturbances in total hip and knee arthroplasty based on an enhanced recovery after surgery concept. *Sleep Breathing*. (2021) 25:1231–7. doi: 10.1007/s11325-020-02213-y
- Dong Y, Peng C. Principled missing data methods for researchers. *SpringerPlus*. (2013) 2:222. doi: 10.1186/2193-1801-2-222
- Haefeli M, Elfering A. Pain assessment. *Eur Spine J*. (2006) 15:S17–24. doi: 10.1007/s00586-005-1044-x
- Zung WWK. A Rating instrument for anxiety disorders. *Psychosomatics*. (1971) 12:371–9. doi: 10.1016/S0033-3182(71)71479-0
- Zung WWK. A self-rating depression scale. *Arch General Psychiatry*. (1965) 12:63–70. doi: 10.1001/archpsyc.1965.01720310065008
- Bellamy N, Buchanan W, Goldsmith C, Campbell J, Stitt L. 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.
- Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B*. (1996) 58:267–88. doi: 10.1111/j.2517-6161.1996.tb02080.x
- Faisal A, Jhanjhi N, Ashraf H, Ray S, Ashfaq F. A comprehensive review of machine learning models: principles, applications, and optimal model selection. *TexRxiv [Preprint]* (2025):doi: 10.36227/techrxiv.174285687.71966152/v1
- Bradley A. The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recogn*. (1997) 30:1145–59. doi: 10.1016/S0031-3203(96)00142-2
- Trifonova O, Lokhov P, Archakov A. [Metabolic profiling of human blood]. *Biomeditsinskaya khimiia*. (2014) 60:281–94. doi: 10.18097/pbmc20146003281
- Cho S, Joo B, Park M, Ahn S, Suh S, Park Y, et al. A radiomics-based model for potentially more accurate identification of subtypes of breast cancer brain metastases. *Yonsei Med J*. (2023) 64:573–80. doi: 10.3349/ymj.2023.0047
- Giuste F, Shi W, Zhu Y, Naren T, Isgut M, Sha Y, et al. Explainable artificial intelligence methods in combating pandemics: a systematic review. *IEEE Rev Biomed Eng*. (2023) 16:5–21. doi: 10.1109/RBME.2022.3185953
- Ozenne B, Subtil F, Maucourt-Boulch D. The precision–recall curve overcame the optimism of the receiver operating characteristic curve in rare diseases. *J Clin Epidemiol*. (2015) 68:855–9. doi: 10.1016/j.jclinepi.2015.02.010
- Okkonen E, Vanhanen H. Family support, living alone, and subjective health of a patient in connection with a coronary artery bypass surgery. *Heart Lung J Cardiopulmonary Acute Care*. (2006) 35:234–44. doi: 10.1016/j.hrtlng.2005.11.002
- Park N, Jang Y, Lee B, Chiriboga D. The relation between living alone and depressive symptoms in older korean americans: do feelings of loneliness mediate? *Aging Mental Health*. (2017) 21:304–12. doi: 10.1080/13607863.2015.1099035
- Siciliani L, Wen J, Gaughan J. Living alone and provider behaviour in public and private hospitals. *J Health Econ*. (2025) 102:103016. doi: 10.1016/j.jhealeco.2025.103016
- Halawi M, Chiu D, Gronbeck C, Savoy L, Williams V, Cote M. Psychological distress independently predicts prolonged hospitalization after primary total hip and knee arthroplasty. *J Arthroplasty*. (2019) 34:1598–601. doi: 10.1016/j.arth.2019.03.063
- Routasalo P, Savikko N, Tilvis R, Strandberg T, Pitkälä K. Social contacts and their relationship to loneliness among aged people – a population-based study. *Gerontology*. (2006) 52:181–7. doi: 10.1159/000091828
- Tong H, Hou W, Liang L, Li T, Liu H, Lee T. Age-related differences of rumination on the loneliness–depression relationship: evidence from a population-representative cohort. *Innov Aging*. (2021) 5:igab034. doi: 10.1093/geroni/igab034
- Widhowati S, Chen C, Chang L, Lee C, Fetzer S. Living alone, loneliness, and depressive symptoms among indonesian older women. *Health Care Women Int*. (2020) 41:984–96. doi: 10.1080/07399332.2020.1797039
- Koo J, Son N, Yoo K. Relationship between the living-alone period and depressive symptoms among the elderly. *Arch Gerontol Geriatr*. (2021) 94:104341. doi: 10.1016/j.archger.2021.104341
- Munir S, Khan S, Nazneen S, Ahmad S. Temporal and seasonal variations of noise pollution in urban zones: a case study in Pakistan. *Environ Sci Pollut Res*. (2021) 28:29581–9. doi: 10.1007/s11356-021-12738-8
- Guha A, Gokhale S. Urban workers' cardiovascular health due to exposure to traffic-originated Pm2.5 and noise pollution in different microenvironments. *Sci Total Environ*. (2023) 859:160268. doi: 10.1016/j.scitotenv.2022.160268
- Linares Arroyo H, Abascal A, Degen T, Aubé M, Espey B, Gyuk G, et al. Monitoring, trends and impacts of light pollution. *Nat Rev Earth Environ*. (2024) 5:417–30. doi: 10.1038/s43017-024-00555-9
- Jones T, Durrant J, Michaelides E, Green M. Melatonin: a possible link between the presence of artificial light at night and reductions in biological fitness. *Philos Trans R Soc B Biol Sci*. (2015) 370:20140122. doi: 10.1098/rstb.2014.0122

50. Blume C, Garbaza C, Spitschan M. Effects of light on human circadian rhythms, sleep and mood. *Somnologie*. (2019) 23:147–56. doi: 10.1007/s11818-019-00215-x
51. Smith Michael G, Cordoza M, Basner M. Environmental noise and effects on sleep: an update to the who systematic review and meta-analysis. *Environ Health Perspect*. (2022) 130:076001. doi: 10.1289/EHP10197
52. Monazzam M, Shamsipour M, Zaredar N, Bayat R. Evaluation of the relationship between psychological distress and sleep problems with annoyance caused by exposure to environmental noise in the adult population of Tehran Metropolitan city. *Iran. J Environ Health Sci Eng*. (2022) 20:1–10. doi: 10.1007/s40201-021-00703-z
53. Xu Y, Zhang J, Tao F, Sun Y. Association between exposure to light at night (Lan) and sleep problems: a systematic review and meta-analysis of observational studies. *Sci. Total Environ*. (2023) 857:159303. doi: 10.1016/j.scitotenv.2022.159303
54. Zhong C, Longcore T, Benbow J, Chung N, Chau K, Wang S, et al. Environmental influences on sleep in the california teachers study cohort. *Am J Epidemiol*. (2022) 191:1532–9. doi: 10.1093/aje/kwab246
55. Liu S, Zhou S, Li Y, Cao L, Lv G, Peng L, et al. Lag analysis of the effect of air pollution on orthopedic postoperative infection in hebei province and Xinjiang Uygur autonomous region. *Sci Rep*. (2025) 15:12919. doi: 10.1038/s41598-025-95550-5
56. Oslock W, Wood L, Sawant A, English N, Jones B, Martin C, et al. Short-term exposure to ambient particulate matter pollution and surgical outcomes. *J Surg Res*. (2025) 307:148–56. doi: 10.1016/j.jss.2025.01.011
57. Griffin S, Ravyts S, Bourchtein E, Ulmer C, Leggett M, Dzierzewski J, et al. Sleep disturbance and pain in U.S. adults over 50: evidence for reciprocal, longitudinal effects. *Sleep Med*. (2021) 86:32–9. doi: 10.1016/j.sleep.2021.08.006
58. Bonvanie I, Oldehinkel A, Rosmalen J, Janssens K. Sleep problems and pain: a longitudinal cohort study in emerging adults. *Pain*. (2016) 157:957–63. doi: 10.1097/j.pain.0000000000000466
59. Chen T, Lee S, Schade M, Saito Y, Chan A, Buxton O. Longitudinal relationship between sleep deficiency and pain symptoms among community-dwelling older adults in Japan and Singapore. *Sleep*. (2019) 42:zsy219. doi: 10.1093/sleep/zsy219
60. Quartana P, Wickwire E, Klick B, Grace E, Smith M. Naturalistic changes in insomnia symptoms and pain in temporomandibular joint disorder: a cross-lagged panel analysis. *Pain*. (2010) 149:325–31. doi: 10.1016/j.pain.2010.02.029
61. Butris N, Tang E, Pivetta B, He D, Saripella A, Yan E, et al. The prevalence and risk factors of sleep disturbances in surgical patients: a systematic review and meta-analysis. *Sleep Med Rev*. (2023) 69:101786. doi: 10.1016/j.smrv.2023.101786
62. Kopel J, Jakubski S, Al-Mekdash M, Berdine G. Distribution of age and apnea-hypopnea index in diagnostic sleep tests in West Texas. *Proceedings*. (2022) 35:15–9. doi: 10.1080/08998280.2021.1966710
63. Hamilton K, Hughes A, Campbell C, Owens M, Pester B, Meints S, et al. Perioperative insomnia trajectories and functional outcomes after total knee arthroplasty. *Pain*. (2023) 164:2769–79. doi: 10.1097/j.pain.0000000000002977