Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Development and validation of an interpretable machine learning model for predicting moderate-to-severe urinary retention risk after mixed hemorrhoidectomy

Lu Ze
Jiangsu Province Hospital of Chinese Medicine, Affiliated Hospital of Nanjing University of Chinese Medicine

Minmin Xu
Jiangsu Province Hospital of Chinese Medicine, Affiliated Hospital of Nanjing University of Chinese Medicine

Long Peng
Jiangsu Province Hospital of Chinese Medicine, Affiliated Hospital of Nanjing University of Chinese Medicine

Dong Chen
Jiangsu Province Hospital of Chinese Medicine, Affiliated Hospital of Nanjing University of Chinese Medicine

Pingping He
hepp@njucm.edu.cn

Jiangsu Province Hospital of Chinese Medicine, Affiliated Hospital of Nanjing University of Chinese Medicine

Article

Keywords:

Posted Date: November 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9891659/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Development and validation of an interpretable machine learning model for predicting moderate-to-severe urinary retention risk after mixed hemorrhoidectomy

Lu Ze [a], Minmin Xu[a] , Long Peng [a], Dong Chen[a*], Pingping He [a*]

Lu Ze and **Minmin Xu** contributed equally to this work

[*]**Corresponding Authors:**

Pingping He (Corresponding Author, hepp@njucm.edu.cn)

Dong Chen (Co-corresponding Author, jsnjcd@163.com)

[a]Jiangsu Province Hospital of Chinese Medicine, Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing, Jiangsu, China

Abstract

Objective: To develop and validate an interpretable machine learning (ML) model for early prediction of moderate-to-severe postoperative urinary retention (POUR) after mixed hemorrhoidectomy. **Methods:**Data from 410 patients (training/internal validation) were retrospectively collected (May 2024–March 2025), and data from 261 patients (external validation) were prospectively collected (April–August 2025). Six ML algorithms (GLM, RF, XGBoost, LASSO, SVM, KNN) were compared using AUC, Brier score, Hosmer-Lemeshow test and DCA. SHAP analysis provided model interpretability. **Results:** Among the six ML models, the LASSO model performed best. This model accurately predicted the risk of moderate-to-severe POUR in patients with mixed hemorrhoidectomy in both the training set (AUC: 0.881, Specificity: 0.979, F1 Score: 0.466) and the external validation set (AUC: 0.739, Specificity: 0.888, F1 Score: 0.333), Brier score=0.150, Hosmer-Lemeshow $P<0.001$, DCA indicated that the LASSO model provided a higher net clinical benefit. SHAP analysis identified the features contributing to the model predictions, 16 variables were included in the final LASSO model: age, glucose, preoperative fasting, insomnia, operative time, constipation, number of incisions, history of urological diseases, urine_white blood cell (U_WBC), analgesic B (Parecoxib), urine_red blood cell (U_RBC), analgesic C (Dezocine), serum calcium, preemptive analgesia, serum potassium, urine glucose.

An online calculator is available at https://mrtools.shinyapps.io/shiny3/. **Conclusion:** An interpretable LASSO-based ML model was successfully developed and externally validated for predicting moderate-to-severe POUR after mixed hemorrhoidectomy. It enables early identification of high-risk patients and supports individualized perioperative management.

Keywords: Mixed hemorrhoidectomy; moderate-to-severe POUR; Machine learning; Interpretable; Predictive model

Introduction

Mixed hemorrhoidectomy is the primary curative treatment for severe hemorrhoidal disease[1]. Despite its therapeutic efficacy, postoperative urinary retention (POUR) remains a highly prevalent complication, with reported incidence rates ranging from 13.7% to 21.9% following anorectal surgery [2][3]. The underlying pathophysiology is multifactorial, involving dense perianal neural innervation, postoperative pain, and reflex sympathetic excitation triggered by rectal distension [4]. Unmanaged POUR significantly increases patient discomfort and elevates the risk of secondary complications, including urinary tract infections and potential long-term bladder dysfunction [5][6][7].

Current clinical practice advocates a graded management approach, escalating from non-pharmacological measures to pharmacologic therapy [8][9] and, ultimately, catheterization for moderate-to-severe POUR [10][11] However, the reliance on indwelling catheterization is suboptimal, over half of catheterized patients experience related complications, with non-infectious adverse events such as urethral injury being notably common [7]. This underscores the imperative to shift from reactive management to proactive, risk-based prevention.

Although several predictive models for POUR following mixed hemorrhoidectomy have been developed, but they exhibit notable limitations. Firstly, most existing studies [2][3][4][5] focus on predicting the binary outcome of “whether POUR occurs”, which offers limited clinical utility. Given that mild POUR is often self-limiting and can be relieved through simple measures such as local heat or auditory stimulation [10], it

seldom impacts patient recovery substantially. Relying solely on the occurrence of POUR for clinical intervention may lead to inefficient resource allocation, potentially resulting in overtreatment of mild cases while under-preventing moderate-to-severe cases, thereby exposing patients to higher risks. Secondly, many current models focus on demographic characteristics and perioperative parameters while overlooking relevant laboratory indicators, such as complete blood count, urine routine analysis, and serum biochemical parameters, despite evidence suggesting that inflammatory markers and blood glucose are associated with the development of POUR [4][12]. The exclusion of these variables may compromise predictive accuracy. Finally, ML techniques have been underutilized in this field, and where applied, models often lack interpretability, failing to fully capture complex variable relationships or provide clinically actionable insights.

To address this gap, this study shifted the predictive focus from the binary occurrence of POUR to a clinically meaningful stratification of its severity. In addition to demographic and perioperative parameters, we integrated routine laboratory indicators, including complete blood count, urine routine analysis, and serum biochemical parameters, with six machine learning (ML) algorithms: generalized linear model (GLM), random forest (RF), extreme gradient boosting (XGB), least absolute shrinkage and selection operator (LASSO), support vector machine (SVM), and K-nearest neighbor (KNN). This approach was used to develop and validate an interpretable ML model for the early and accurate prediction of moderate-to-severe POUR risk in patients undergoing mixed hemorrhoidectomy. Using the SHAP method, we clarified feature importance and provided explanations for individual predictions. Finally, we deployed the prediction model as a Shiny application-based web platform, with the ultimate goal of facilitating stratified prevention, optimizing clinical decision-making, and improving both patient outcomes and healthcare resource efficiency.

Methods

Study design and patient selection

We collected the data of patients who underwent the hemorrhoidectomy at the Department of Colorectal Surgery of a traditional Chinese medicine hospital in Nanjing between May 2024 and August 2025. The inclusion criteria were as follows: (1) Diagnosis of mixed hemorrhoids and candidate for hemorrhoidectomy under subarachnoid block anesthesia; (2) Age between 18 and 80 years; (3) With complete medical records. Exclusion criteria: (1) Preexisting neurogenic bladder; (2) Long-term indwelling urinary catheterization; (3) Concomitant severe cardiac, hepatic, or renal dysfunction; (4) Current pregnancy or lactation.

Ethics approval for this study was granted by the Ethics Committee of the Affiliated Hospital of Nanjing University of Chinese Medicine (Jiangsu Provincial Hospital of Chinese Medicine) (Approval No. 2024NL-325-02). For the retrospective component, data were extracted from pre-existing medical records and anonymised, which allowed the requirement for informed consent to be waived. For the prospective component, written informed consent was obtained from all participating patients. The study was performed in strict compliance with the TRIPOD guidelines. All methods were performed in accordance with the relevant guidelines and regulations.

Sample size calculation

This study employed a two-phase design. First, a prediction model was developed and internally validated using a retrospectively assembled cohort. Subsequently, external validation was performed in a prospectively collected, single-center cohort. The retrospective cohort consisted of 410 consecutive patients undergoing mix hemorrhoidectomy and served as the model development set. An independent prospective cohort of 261 consecutive patients was enrolled specifically to evaluate the model performance in new data.

For logistic regression prediction models, a key metric for sample size adequacy is the

events-per-variable (EPV) ratio. The reported incidence of POUR after anorectal surgery in the literature ranges from 13.7% to 21.9% [3][4][13]. Based on preliminary research and clinical expertise, eight candidate predictor variables were pre-specified. In the final retrospective development cohort (n=410), 71 moderate-to-severe POUR events were observed. The actual EPV during model development was therefore 8.9. Although this value falls slightly below the methodologically recommended ideal of EPV ≥10, the following strategies were implemented to enhance model reliability: (1) regularization techniques such as LASSO regression were applied for variable selection to mitigate overfitting; and (2) bootstrap resampling was used during model development for internal validation, yielding stable and minimally biased performance estimates.

The sample size of the prospective validation cohort (n=261) was determined by the number of consecutive patients available within the predefined study period. This cohort provided 46 moderate-to-severe POUR events. Its primary aim was to offer a preliminary assessment of the model discriminative ability and calibration in an independent dataset, rather than to reestimate model parameters.

Clinical features

We established a multi-dimensional framework for variables collection, encompassing three complementary domains: baseline demographics and behavioral variables, perioperative variables, and laboratory indicators (within 24 hours of the perioperative period). This inclusive strategy, guided by clinical expertise and existing literature [3][14]was designed to capture a comprehensive informational landscape and mitigate preselection bias. The data for the retrospective cohort were extracted through structured queries from the hospitalization electronic medical records (EMRs). For the prospective cohort, data were collected in real time using pre-designed electronic case report forms upon patient admission to ensure data quality and completeness. Data collection for both cohorts followed identical variable definitions and standards.

Baseline demographics and behavioral variables included age, body mass index (BMI), gender (0=male, 1=female), insomnia (0=No, 1=Yes), constipation (0=No, 1=Yes), history of urological diseases (0=No, 1=Yes), diabetes (0=No, 1=Yes).

Perioperative variables included preoperative fasting, operative time, number of incisions, total intravenous fluid volume, antiemetic (0=No, 1=Yes), analgesic type (A=Flurbiprofen Axetil Injection, B=Parecoxib Sodium, C=Dezocine), supplemental analgesia (0=No, 1=Yes), preemptive analgesia (0=No, 1=Yes).

Laboratory indicators included urine_white blood cell (U_WBC), urine_red blood cell (U_RBC), blood_white blood cell (B_WBC), blood_red blood cell (B_RBC), urine glucose (0=negative, 1=positive), U_bacteria (0=negative, 1=positive), and U_blood (0=negative, 1=positive). hemoglobin (Hb), neutrophil percentage (Neutrophil_Pct), neutrophil absolute (Neutrophil_Abs), calcium (Ca), creatinine (Cr), serum potassium (K), serum sodium (Na), glucose, and total protein (TP).

Assessment of study outcome

The primary outcome was moderate-to-severe POUR, which was assessed within the first 24 hours after hemorrhoidectomy. Severity was defined as follows: (1) Mild POUR: initial postoperative difficulty in urination that resolved spontaneously or with non-pharmacological measures (Chinese herbal heat therapy or moxibustion); (2) Moderate POUR: initial difficulty unresponsive to non-pharmacological interventions but effectively relieved by pharmacological treatment (intramuscular neostigmine or tamsulosin sustained-release capsules orally); (3) Severe POUR: initial difficulty refractory to all the above measures, or a post-void residual bladder volume >200 mL, necessitating indwelling catheterization based on patient preference and clinical judgment. Classification criteria for moderate-to-severe POUR were consistently documented in medical orders, ensuring data reliability and objectivity. For model development, mild POUR cases were combined with no-POUR cases (coded as 0), and moderate and severe POUR cases were combined (coded as 1), creating a binary outcome variable.

Data preprocessing

To more intuitively visualize variable distributions, this study employed the following graphical methods: For continuous variables, a matrix of box plots, histograms, kernel density estimation curves, and violin plots was generated to comprehensively visualize their distribution patterns, central tendencies, and potential outliers. Additionally,

restricted cubic spline (RCS) plots were used to depict potential non-linear trends **(Supplementary Fig.1**). For categorical variables, frequency bar charts with percentage annotations were created to display the proportional composition of different categories (**Supplementary Fig.2**). This step primarily served to understand data characteristics, identify obvious outliers and unreasonable distributions, and provide a basis for subsequent modeling and variable processing.

Missing data handling and collinearity assessment

Within the modeling cohort, the 'mice' package was used for multiple imputation of missing variables (missing rate < 20%), configured with an imputation count of m=5 and a maximum iteration number of maxit=10, thereby generating 5 complete imputed datasets. Following the imputation of missing values, all candidate predictor variables were included in a complete multivariable logistic regression model to assess multicollinearity. This assessment was conducted by calculating the variance inflation factor (VIF) for each variable. A VIF exceeding 4 was considered indicative of high multicollinearity, leading to the removal of variables including B_WBC, Neutrophil_Pct, and Neutrophil_Abs, to reduce information redundancy and enhance model stability (**Supplementary Table 1**). Upon completion of 'mice' imputation and collinearity treatment, the imputed complete data within the retrospective cohort were randomly partitioned into training and internal validation sets in a 7:3 ratio.

Selection of variables

To identify key predictors while balancing model stability and interpretability, this study employed a comprehensive approach utilizing multiple feature selection methods based on different algorithms to evaluate all candidate predictor variables, the features selected through this process were detailed in **Supplementary Table 2.**

Model development and performance evaluation

This study developed six ML models to predict the risk of moderate-to-severe POUR in patients undergoing mixed hemorrhoidectomy. Each model was independently trained using the predictor set selected by its corresponding feature selection method. To obtain more stable probability estimates and mitigate the influence of sampling variability, this study applied a bootstrap resampling strategy with 100 repetitions to all

six models within the training set. This process yielded six sets of bootstrap-smoothed predicted probabilities, which were subsequently used for performance evaluation in the training set **(Supplementary Table 3).**

Model performance was assessed using several commonly used evaluation indicators, including the AUC, sensitivity, specificity, accuracy, F1 score, and Brier score. The Hosmer-Lemeshow test was used to examine the goodness-of-fit between predicted probabilities and observed outcomes. Decision curve analysis (DCA) was performed to evaluate the clinical net benefit of the models across various threshold probabilities. The optimal prediction model was selected based on its comprehensive performance across the aforementioned metrics in both the training and validation sets.

Model explanation

Explaining ML models is challenging. The SHAP method is a game theory-based technique that ranks the importance of input features and explains the results of the prediction model, overcoming the “black box” problem. The SHAP method provides explanations by calculating the contribution of each feature to the prediction results, thus increasing the model transparency and interpretability.

Network calculator

To facilitate model application in a clinical setting, the final prediction model was integrated into a Shiny application-based web platform. When the values of the relevant features in the final model are provided, this application returns the probability of moderate-to-severe POUR in patients with mixed hemorrhoidectomy.

Statistics

These ML models were developed using R (Version 4.3.1) and the “caret” (Version: 6.0.94) package. “caret” is a comprehensive package that provides a unified interface for various ML algorithms. The discrimination performance was evaluated using the ROC curve analysis, and the AUC and its biascorrected 95% confidence interval (CI) using 100-fold bootstrap were reported [15]. The Brier score (ranging from 0 to 1) was used to calculate the difference between the estimated risk and the observed risk, with a value closer to 0 indicating better calibration, thus assessing model calibration. In addition, the calibration performance of the clinical prediction models was evaluated

by the Hosmer-Lemeshow test, with P-values higher than 0.05 usually indicating a good fit between the model and the actual data [16][17]. DCA was conducted to show the net benefit of using a model at different thresholds to assess the clinical value of the model [18].

Continuous variables with normal distribution were presented as mean ± standard deviation and were compared with the t-test. Continuous variables with skewed distributions were presented as medians with interquartile ranges and compared with the Mann-Whitney U test or the Kruskal-Wallis H test. Categorical variables were presented as numbers with percentages and compared with the chi-square test. Independent risk factors were determined by univariable and multivariable logistic regression analyses. A two-tailed P-value < 0.05 was considered statistically significant.

Results

Baseline clinical information

In the training set of 410 patients, 44 cases developed moderate POUR and 27 cases severe POUR, resulting in a total of 71 patients (17.3%) with moderate-to-severe POUR. Similarly, in the external validation set of 261 patients, 31 and 15 cases developed moderate and severe POUR, respectively, totaling 46 patients (17.6%). The distribution of most variables was comparable between the training and internal validation sets, with no significant differences observed for the majority (most P-values were higher than 0.05), indicating good homogeneity after randomization, details are shown in **Table 1.** Furthermore, significant differences were found between the no-POUR and moderate-to-severe POUR groups regarding the following variables (P<0.05): number of incisions, glucose, insomnia, history of urological diseases, analgesia type, and preemptive analgesia (**Table 2**). The study design is detailed in **Fig.1.**

Independent risk factors

This study explored the independent risk factors for moderate-to-severe POUR following hemorrhoidectomy. Using univariate logistic regression analysis, 6 potential

risk factors associated with moderate-to-severe POUR (P<0.05) were identified (**Table 3**). After multivariate logistic regression analysis, 6 factors independently associated with the risk of moderate-to-severe POUR were finally identified, namely, history of urological diseases (OR=3.949), insomnia (OR=3.073), elevated blood glucose (OR=2.94), number of surgical incisions (OR=1.68), longer fasting hours (OR=0.799), and preemptive analgesia (OR=0.035). Among these, longer fasting hours and preemptive analgesia were protective factors (OR<1), while the remaining variables were risk factors (**Table 4**).

Model development and performance comparison

In the training set, the RF and KNN models achieved perfect classification (AUC = 1), yet their performance plummeted in internal validation (RF AUC dropped to 0.702, KNN AUC to 0.577), suggesting severe overfitting. Hence, they were not selected as the final models. For discrimination, XGB performed best (AUC=0.991, Specificity=1, F1=0.563), followed by LASSO (AUC=0.881, Specificity=0.979, F1=0.466) (**Fig.2**). For calibration, XGB had the lowest Brier score (0.069) but failed the Hosmer–Lemeshow test (P<0.001), while LASSO had a higher Brier score (0.100) but passed calibration (P=0.283) **(Table 5)**. DCA suggested that XGB and SVM offered higher net benefit across most thresholds, whereas LASSO showed more stable performance (**Fig.3**).

In the internal validation set, regarding discrimination (**Fig.4)**, the GLM model performed best (AUC=0.805, Specificity=0.971, F1=0.467), followed by the LASSO model (AUC=0.761, Specificity=0.971, F1=0.414). In terms of calibration (**Table 6**), the GLM model had the lowest Brier score (Brier-score=0.109, H-L test =0.088), followed by the LASSO model (Brier-score=0.116, H-L test =0.014). DCA suggested that GLM and LASSO model offered higher net benefit across most thresholds (**Fig.5**). Therefore, LASSO was selected as the final model due to its stable performance across datasets, good calibration, simple interpretable structure, and optimal balance between predictive power and clinical applicability for moderate-to-severe POUR risk prediction.

External validation of the LASSO model

For external validation, LASSO model was independently evaluated on an external cohort from April 2025 to August 2025 (n=261). The performance in the external validation set demonstrated that the LASSO model maintained acceptable discriminatory power, with an AUC of 0.739 (**Supplementary Table 4**), but required recalibration (Brier score=0.150; Hosmer-Lemeshow P<0.001), refer to **Supplementary Table 5**. DCA indicated that the LASSO model provided a higher net clinical benefit (**Fig. 6**). In summary, external validation confirmed the LASSO model's generalizable despite the observed decline in purely statistical metrics, performance, supporting its potential for clinical use in similar settings.

Correlation Heatmap analysis of LASSO model variables

Spearman correlation analysis was employed to visualize the degree of correlation between various SHAP features and the model predictions, aiding in the identification of variable relationships and potential collinearity. This is a symmetric heatmap displaying the pairwise correlation coefficients among the 15 features selected by the LASSO model. The axes (X-axis and Y-axis) typically share the same ordered list of features. “dark blue” indicates a strong negative correlation (close to -1), “white or light colors” indicate a weak or negligible correlation (close to 0), and “dark red” indicates a strong positive correlation (close to +1). The cell values display the specific correlation coefficients. For example, the correlation coefficient between “preemptive analgesia” and the outcome (moderate-to-severe POUR) is -0.27, indicating a negative correlation, which suggests that preemptive analgesia can reduce the incidence of moderate-to-severe POUR. The correlation coefficient between “urinary red blood cells” and “urinary white blood cells” is +0.47, indicating a positive correlation between the two variables (**Fig. 7**).

Model explanation

Since it is difficult for clinicians to accept prediction models that are not directly explainable or not explainable at all, this study performed SHAP interpretation analysis based on the screened LASSO model to assess the importance of key predictive variables and their impact on model output. As shown in the SHAP waterfall plot (**Fig.**

8), the contribution of features to the model was evaluated using the mean absolute SHAP values, which were displayed in descending order: U-RBC, history of urological diseases, analgesic type (C/A), insomnia, and operative time were the five most important features in increasing the predicted probability, while number of incisions, U-WBC, preemptive analgesia, U-glucose, and Ca were the five most important features whose numerical changes were associated with a decrease in the predicted probability.

In addition, the SHAP dependence plot helps to understand how individual features affect the output of the prediction model. It illustrates the trend of a feature impact on predicted risk as its value increases. An upward curve indicates an increased risk, while a downward curve indicates a decreased risk (**Supplementary Fig. 3**). This plot compared the distributions of SHAP values across the actual values of these 16 features, where features with SHAP values greater than zero contribute positively to the model's prediction, that is, these features indicate a higher risk of moderate-to-severe POUR. Overall, elevated blood glucose (values >5 mmol/L), insomnia, a higher number of surgical incisions (>3), prolonged operative time (>30 minutes), constipation, and the presence of U_RBC all increase the risk of moderate-to-severe POUR. In contrast, preemptive analgesia, a longer fasting duration (>19 hours), U_WBC, the type of analgesic administered, calcium levels (>2.25 mmol/L), older age (>45 years), and higher serum potassium levels (>3.9 mmol/L) were associated with a reduced risk of moderate-to-severe POUR.

Implementation of the web calculator

As shown in **Fig. 9,** the final prediction model was integrated into a web application for use in clinical settings. By inputting the actual values of the 16 features required for the model, the application can automatically predict the risk of moderate-to-severe POUR in individual patients with mixed hemorrhoidectomy. The web application can be accessed online at the following link: https://mrtools.shinyapps.io/shiny3/.

Discussion

While previous studies focused on binary POUR occurrence [19][20][21][22], their

dichotomous approach fails to distinguish mild from moderate-to-severe cases. To address this gap, we present, to our knowledge, the first ML model specifically designed to stratify POUR by severity. Focusing on patients undergoing mixed hemorrhoidectomy, we developed and compared six ML algorithms to predict the risk of moderate-to-severe POUR and identified a set of key predictive risk factors.

Among the six ML models evaluated, the LASSO model demonstrated favorable performance in terms of accuracy, calibration, reclassification and overall discriminatory ability, net benefit, and interpretability. Specifically, the LASSO model effectively balanced predictive power with model simplicity, avoided overfitting, and provided stable performance across both internal and external validation. LASSO regression was utilized for variable selection, which, by penalizing the regression, effectively reduces unnecessary predictive variables and addresses issues of collinearity, particularly suitable for high-dimensional datasets [23][24][25]. In this study, we developed a final prediction model using LASSO regression that incorporates 16 features. These features can be easily obtained and assessed during the hospitalization of patients undergoing mixed hemorrhoidectomy, making the model a promising tool for effectively predicting the risk of moderate-to-severe POUR after surgery.

Multivariable logistic regression identified independent predictors: history of urological diseases (OR=3.949), insomnia (OR=3.073), elevated glucose (OR=2.94), number of incisions (OR=1.68), longer fasting (OR=0.799), and preemptive analgesia (OR=0.035).

History of urological diseases was the strongest risk factor (OR=3.949), consistent with prior studies [2][3], we recommend incorporating patients with history of urological diseases, into high-risk clinical pathways for POUR management. Routine screening for urinary symptoms is recommended for all patients scheduled for anorectal surgery, in high-risk patients, consideration should be given to urodynamic evaluation or prophylactic use of oral tamsulosin [26]. Insomnia (OR=3.073) significantly increased risk, possibly via autonomic dysfunction and exacerbated pain [27][28][29]. Notably, approximately 29.8% of patients in this study had insomnia, a prevalence rate

consistent with the general population insomnia rate, approximately 30% [30], it highlighting the clinical significance of this finding. Based on these findings, we recommend incorporating sleep quality into preoperative assessment systems. For patients with sleep disorders, cognitive behavioral therapy or traditional Chinese medicine interventions should be considered. Elevated glucose (OR=2.94) was associated with moderate-to-severe POUR. Hyperglycemia impairs autonomic nerve and detrusor function, exacerbates inflammation, and affects pelvic nerves [31][32]. Thus, rational control of blood glucose is of significant importance in preventing the severity of POUR, guiding clinical healthcare providers to pay close attention to preoperative blood glucose levels and to actively implement preventive measures for diabetic patients. Increased incisions (OR=1.68) reflect greater tissue trauma, promoting edema and pain that inhibit detrusor function, more than three incisions significantly raise POUR risk [2][3]. Preemptive wrist-ankle acupuncture (WAA) was protective (OR=0.035). WAA is a form of acupuncture therapy that has been used clinically for many years, preemptive analgesia using acupuncture can prevents central sensitization, reduces neuroendocrine stress and inflammation, alleviates pelvic floor spasms, and decreases postoperative analgesic needs, avoiding bladder contractility inhibition [33]. This finding supports the routine implementation of preemptive analgesia during the perioperative period of mixed hemorrhoids to optimize urinary outcomes, particularly those TCM-based analgesic methods that are side-effect-free, safe, and cost-effective. Longer fasting (OR=0.799) was associated with lower POUR risk. This does not advocate prolonged fasting but suggests that within ERAS principles, earlier return to the ward warrants closer voiding monitoring.

The final LASSO model included 16 variables: age, glucose, preoperative fasting, insomnia, operative time, constipation, number of incisions, history of urological diseases, U_WBC, analgesic B (Parecoxib), U_RBC, analgesic C (Dezocine), Ca, K, preemptive analgesia, and urine glucose. The correlation heatmap, SHAP waterfall plot, and dependence plots collectively revealed that the associations of preemptive analgesia, elevated blood glucose, insomnia, and a history of urological diseases with the outcome of moderate-to-severe POUR were consistent with the results from the

aforementioned multivariate analysis. Furthermore, the dependence plots indicated that patients who received “preemptive analgesia” or were administered “non-opioid analgesics” were less likely to develop moderate-to-severe POUR. Conversely, patients with “higher blood glucose levels, more than 3 surgical incisions, operation times exceeding 30 minutes, constipation, insomnia, or elevated U_RBC counts” were more prone to develop moderate-to-severe POUR.

ML models are often criticized as “black box” due to their limited interpretability, which may hinder clinical adoption [34]. A key strength of our study is the application of SHAP analysis to enhance model transparency. SHAP not only quantifies feature importance but also explains individualized predictions, thereby supporting clinical reasoning. Additionally, we deployed the final LASSO model via an interactive Shiny web application, offering a user-friendly tool for both physicians and patients. This calculator can serve as a mobile decision-making assistant tool, helping nurses or junior doctors identify high-risk patients early after hemorrhoidectomy, thereby enabling them to take preventive measures in advance. Another strength is the systematic comparison of six ML algorithms for predicting moderate-to-severe POUR after mixed hemorrhoidectomy, with external validation confirming the LASSO model’s robust predictive performance. Moreover, all predictors included are routinely collected during hospitalization, facilitating model implementation in everyday practice.

Several limitations should be acknowledged. First,This might be due to the significant differences between the training set and the external validation set in aspects such as "analgesic type" ($p < 0.001$ in Table 1, almost all in analgesic type A (A=Flurbiprofen Axetil Injection) in the external set), which affected the absolute accuracy of the model's probability predictions. It is recommended that in the future, the model coefficients should be updated through multi-center data. Second, this is a single center study, although external validation was performed, generalizability to other centers requires further confirmation. Third, no established criterion determines the optimal sample size for prediction modeling, yet the satisfactory internal and external

validation results suggest our sample was adequate for exploring moderate-to-severe POUR risk in this population. Fourth, in the interest of generalizability, we included only routine clinical variables; novel urodynamic parameters such as postvoid residual volume (PRV), which may improve diagnostic accuracy and model performance, were not incorporated. Nonetheless, these limitations do not diminish the overall favorable performance of our final prediction model.

Conclusions

In summary, we successfully developed an interpretable ML model to predict the risk of moderate-to-severe POUR in patients after mixed hemorrhoidectomy, based on clinical data that can be readily extracted from EMR. The final LASSO model demonstrated robust predictive performance in both internal and external validation. In the future, further prospective RCTs are warranted to clarify whether personalized treatment plans and follow-up strategies formulated based on the final prediction model can effectively reduce the incidence of moderate-to-severe POUR following mixed hemorrhoidectomy.

The authors did not use generative AI or AI-assisted technologies in the development of this manuscript.

Table 1 Comparison of characteristics between the training and validation sets

Characteristic	level	Training set (n=287)	Internal validation set (n=123)	External validation set (n=261)	P
age (mean (SD))		43.99 (12.83)	44.22 (12.49)	44.80 (13.77)	0.765
BMI (mean (SD))		23.35 (3.17)	23.39 (3.64)	23.36 (3.06)	0.992
preoperative fasting (mean (SD))		18.77 (2.20)	19.03 (2.43)	18.75 (2.57)	0.529
operative time (mean (SD))		32.10 (11.62)	34.26 (13.36)	32.69 (13.20)	0.28
number of incisions (mean (SD))		3.47 (1.01)	3.48 (0.98)	3.41 (0.96)	0.688
total intravenous fluid volume (mean (SD))		1190.31 (229.22)	1206.07 (190.62)	1062.07 (183.18)	<0.001
U_WBC (mean (SD))		38.24 (109.83)	36.05 (78.31)	32.44 (129.49)	0.835
U_RBC (mean (SD))		18.28 (44.96)	16.76 (31.04)	16.19 (95.15)	0.934
B_RBC (mean (SD))		4.57 (0.52)	4.54 (0.55)	4.53 (0.56)	0.736
Hb (mean (SD))		134.45 (18.13)	133.17 (19.50)	131.25 (20.70)	0.156
Ca (mean (SD))		2.25 (0.11)	2.23 (0.08)	2.25 (0.11)	0.29
Cr (mean (SD))		65.94 (14.93)	63.70 (14.39)	64.63 (14.98)	0.325
K (mean (SD))		3.95 (0.30)	3.88 (0.28)	3.91 (0.29)	0.085
Na (mean (SD))		138.72 (1.99)	138.79 (1.76)	138.99 (1.85)	0.236
Glu (mean (SD))		4.69 (0.73)	4.59 (0.66)	4.87 (0.81)	0.001
TP (mean (SD))		67.32 (5.47)	67.11 (5.60)	67.72 (5.42)	0.531
gender (%)	0	113 (39.4)	46 (37.4)	105 (40.2)	0.869
	1	174 (60.6)	77 (62.6)	156 (59.8)	
insomnia (%)	0	201 (70.0)	88 (71.5)	186 (71.3)	0.932
	1	86 (30.0)	35 (28.5)	75 (28.7)	
constipation (%)	0	136 (47.4)	51 (41.5)	136 (52.1)	0.142
	1	151 (52.6)	72 (58.5)	125 (47.9)	

Characteristic	level	Training set (n=287)	Internal validation set (n=123)	External validation set (n=261)	P
History_UD (%)	0	242 (84.3)	110 (89.4)	218 (83.5)	0.296
	1	45 (15.7)	13 (10.6)	43 (16.5)	
diabetes (%)	0	280 (97.6)	122 (99.2)	249 (95.4)	0.098
	1	7 (2.4)	1 (0.8)	12 (4.6)	
antiemetic (%)	0	80 (27.9)	42 (34.1)	94 (36.0)	0.11
	1	207 (72.1)	81 (65.9)	167 (64.0)	
analgesic type (%)	A	195 (67.9)	80 (65.0)	252 (96.6)	<0.001
	B	83 (28.9)	40 (32.5)	0 (0.0)	
	C	9 (3.1)	3 (2.4)	9 (3.4)	
supplemental analgesia (%)	0	241 (84.0)	106 (86.2)	209 (80.1)	0.269
	1	46 (16.0)	17 (13.8)	52 (19.9)	
preemptive analgesia (%)	0	191 (66.6)	87 (70.7)	194 (74.3)	0.137
	1	96 (33.4)	36 (29.3)	67 (25.7)	
urine glucose (%)	0	256 (89.2)	115 (93.5)	245 (93.9)	0.103
	1	31 (10.8)	8 (6.5)	16 (6.1)	
U_Bacteria (%)	0	266 (92.7)	110 (89.4)	248 (95.0)	0.13
	1	21 (7.3)	13 (10.6)	13 (5.0)	
U_Blood (%)	0	197 (68.6)	73 (59.3)	203 (77.8)	0.001
	1	90 (31.4)	50 (40.7)	58 (22.2)	
outcome	0	236 (82.2)	103 (83.7)	215 (82.4)	0.929
	1	51 (17.8)	20 (16.3)	46 (17.6)	
outcome classification	no	158 (55.1)	69 (56.1)	133 (51.0)	0.836
	mild	78 (27.2)	34 (27.6)	82 (31.4)	
	moderate	30 (10.5)	14 (11.4)	31 (11.9)	
	severe	21 (7.3)	6 (4.9)	15 (5.7)	

Abbr. body mass index (BMI), urine_white blood cell (U_WBC), urine _red blood cell (U_RBC), blood_red blood cell (B_RBC), hemoglobin (Hb), calcium (Ca), creatinine (Cr), serum potassium (K), serum sodium (Na), blood glucose (Glu), total protein (TP), history of urological diseases (History_UD), analgesic type (A=Flurbiprofen Axetil Injection, B=Parecoxib Sodium, C=Dezocine)

455 **Table 2 Comparison of characteristics between patients with and without moderate-to-**

456 **severe POUR in the training set（n=287）**

	level	no-POUR	moderate-to-severe POUR	p	SMD
n		236	51		
age (mean (SD))		43.82 (12.96)	44.80 (12.30)	0.62	0.078
BMI (mean (SD))		23.32 (3.21)	23.48 (3.03)	0.736	0.053
preoperative fasting (mean (SD))		18.86 (2.24)	18.36 (1.97)	0.138	0.239
operative time (mean (SD))		31.76 (11.42)	33.67 (12.51)	0.288	0.159
number of incisions (mean (SD))		3.41 (0.98)	3.75 (1.09)	0.031	0.322
total intravenous fluid volume (mean (SD))		1189.03 (235.71)	1196.24 (198.38)	0.839	0.033
U_WBC (mean (SD))		43.46 (120.25)	14.10 (17.49)	0.083	0.342
U_RBC (mean (SD))		17.40 (41.05)	22.31 (60.20)	0.48	0.095
B_RBC (mean (SD))		4.57 (0.52)	4.58 (0.48)	0.877	0.025
Hb (mean (SD))		134.58 (18.49)	133.82 (16.53)	0.787	0.043
Ca (mean (SD))		2.25 (0.11)	2.24 (0.09)	0.548	0.097
Cr (mean (SD))		66.09 (15.03)	65.22 (14.59)	0.705	0.059
K (mean (SD))		3.95 (0.31)	3.92 (0.25)	0.534	0.102
Na (mean (SD))		138.71 (1.91)	138.77 (2.31)	0.861	0.025
Glu (mean (SD))		4.62 (0.64)	5.03 (0.98)	<0.001	0.495
TP (mean (SD))		67.18 (5.64)	67.96 (4.58)	0.354	0.153

gender (%)	0	91 (38.6)	22 (43.1)	0.654	0.093
	1	145 (61.4)	29 (56.9)		
insomnia (%)	0	173 (73.3)	28 (54.9)	0.015	0.391
	1	63 (26.7)	23 (45.1)		
constipation (%)	0	117 (49.6)	19 (37.3)	0.149	0.251
	1	119 (50.4)	32 (62.7)		
History_UD (%)	0	207 (87.7)	35 (68.6)	0.001	0.475
	1	29 (12.3)	16 (31.4)		
diabetes (%)	0	231 (97.9)	49 (96.1)	0.798	0.105
	1	5 (2.1)	2 (3.9)		
antiemetic (%)	0	66 (28.0)	14 (27.5)	1	0.012
	1	170 (72.0)	37 (72.5)		
analgesic type (%)	A	158 (66.9)	37 (72.5)	0.04	0.353
	B	73 (30.9)	10 (19.6)		
	C	5 (2.1)	4 (7.8)		
supplemental analgesia (%)	0	200 (84.7)	41 (80.4)	0.577	0.115
	1	36 (15.3)	10 (19.6)		
preemptive analgesia (%)	0	143 (60.6)	48 (94.1)	<0.001	0.874
	1	93 (39.4)	3 (5.9)		
urine glucose (%)	0	209 (88.6)	47 (92.2)	0.616	0.122
	1	27 (11.4)	4 (7.8)		
U_Bacteria (%)	0	219 (92.8)	47 (92.2)	1	0.024
	1	17 (7.2)	4 (7.8)		
U_Blood (%)	0	161 (68.2)	36 (70.6)	0.87	0.051
	1	75 (31.8)	15 (29.4)		

Abbr. body mass index (BMI), urine_white blood cell (U_WBC), urine _red blood cell (U_RBC), blood_red blood cell (B_RBC), hemoglobin (Hb), calcium (Ca), creatinine (Cr), serum potassium (K), serum sodium (Na), blood glucose (Glu), total protein (TP), history of

460 urological diseases (History_UD) ,analgesic type (A=Flurbiprofen Axetil Injection,
461 B=Parecoxib Sodium, C=Dezocine)

462 **Table 3 univariate analysis**
463

factor	OR(95%CI)	p.value
age	1.006 (0.982, 1.03)	0.618
BMI	1.017 (0.923, 1.117)	0.735
preoperative fasting	0.899 (0.778, 1.033)	0.139
operative time	1.014 (0.988, 1.04)	0.288
number of incisions	1.362 (1.021, 1.813)	0.034
total intravenous fluid volume	1 (0.999, 1.001)	0.839
U_WBC	0.987 (0.971, 0.997)	0.046
U_RBC	1.002 (0.995, 1.008)	0.484
B_RBC	1.048 (0.583, 1.9)	0.876
Hb	0.998 (0.982, 1.015)	0.787
Ca	0.414 (0.022, 7.134)	0.547
Cr	0.996 (0.975, 1.016)	0.704
K	0.722 (0.256, 1.994)	0.533
Na	1.014 (0.871, 1.184)	0.86
Glu	2.033 (1.367, 3.088)	0.001
TP	1.027 (0.971, 1.086)	0.353
gender-1	0.827 (0.449, 1.54)	0.544
insomnia-1	2.256 (1.204, 4.203)	0.01
constipation-1	1.656 (0.896, 3.131)	0.112
History_UD-1	3.263 (1.586, 6.583)	0.001

factor	OR(95%CI)	p.value
diabetes-1	1.886 (0.265, 9.03)	0.456
antiemetic-1	1.026 (0.531, 2.076)	0.941
analgesic typeB	0.585 (0.263, 1.2)	0.162
analgesic typeC	3.416 (0.812, 13.523)	0.077
supplemental analgesia-1	1.355 (0.597, 2.864)	0.443
preemptive analgesia-1	0.096 (0.023, 0.272)	<0.001
U_Glucose-1	0.659 (0.188, 1.785)	0.456
U_Bacteria-1	1.096 (0.305, 3.124)	0.874
U_Blood-1	0.894 (0.45, 1.706)	0.741

Abbr. body mass index (BMI), urine_white blood cell (U_WBC), urine _red blood cell (U_RBC), blood_red blood cell (B_RBC), hemoglobin (Hb), calcium (Ca), creatinine (Cr), serum potassium (K), serum sodium (Na), blood glucose (Glu), total protein (TP), history of urological diseases (History_UD) ,analgesic type (A=Flurbiprofen Axetil Injection, B=Parecoxib Sodium, C=Dezocine)

Table 4 multivariate logistic regression analysis

factor	OR(95%CI)	p.value
age	0.977 (0.94, 1.014)	0.223
BMI	0.956 (0.826, 1.103)	0.545
preoperative fasting	0.799 (0.66, 0.955)	0.017
operative time	1.025 (0.985, 1.067)	0.232
number of incisions	1.68 (1.088, 2.698)	0.024
total intravenous fluid volume	0.999 (0.997, 1.001)	0.505
U_WBC	0.978 (0.952, 0.995)	0.058
U_RBC	1.005 (0.994, 1.019)	0.423
B_RBC	2.029 (0.567, 7.307)	0.273

factor	OR(95%CI)	p.value
Hb	0.988 (0.95, 1.027)	0.532
Ca	0.051 (0, 8.713)	0.264
Cr	0.976 (0.933, 1.017)	0.259
K	0.634 (0.136, 2.857)	0.556
Na	1.078 (0.866, 1.36)	0.509
Glu	2.94 (1.55, 5.953)	0.002
TP	1 (0.9, 1.113)	0.998
gender-1	0.578 (0.143, 2.258)	0.434
insomnia-1	3.073 (1.298, 7.495)	0.011
constipation-1	1.78 (0.742, 4.379)	0.2
History_UD-1	3.949 (1.485, 10.807)	0.006
diabetes-1	4.197 (0.202, 61.91)	0.321
antiemetic-1	0.986 (0.386, 2.604)	0.977
analgesic type B	0.359 (0.11, 1.074)	0.076
analgesic type C	5.81 (0.828, 43.353)	0.076
supplemental analgesia-1	0.816 (0.267, 2.345)	0.712
preemptive analgesia-1	0.035 (0.005, 0.14)	<0.001
U_Glucose-1	0.337 (0.077, 1.211)	0.115
U_Bacteria-1	1.349 (0.263, 6.332)	0.706
U_Blood-1	1.723 (0.587, 5.001)	0.315

Abbr. body mass index (BMI), urine_white blood cell (U_WBC), urine _red blood cell (U_RBC), blood_red blood cell (B_RBC), hemoglobin (Hb), calcium (Ca), creatinine (Cr), serum potassium (K), serum sodium (Na), blood glucose (Glu), total protein (TP), history of urological diseases (History_UD) ,analgesic type (A=Flurbiprofen Axetil Injection, B=Parecoxib Sodium, C=Dezocine)

Table 5 Assessment of model discrimination and calibration (Training set)

Model	AUC	Sensitivity	Specificity	Accuracy	F1	Brier-score	H-L test (P value)
GLM	0.859	0.235	0.949	0.822	0.32	0.109	0.488
RF	1	0.961	1	0.993	0.98	0.038	<0.001
XGB	0.991	0.392	1	0.892	0.563	0.069	<0.001
LASSO	0.881	0.333	0.979	0.864	0.466	0.1	0.283
SVM	0.849	0.137	1	0.847	0.241	0.116	<0.001
KNN	1	1	1	1	1	0	NA

Table 6 Assessment of model discrimination and calibration (Internal validation set)

Model	AUC	Sensitivity	Specificity	Accuracy	F1	Brier-score	H-L test (P value)
GLM	0.805	0.350	0.971	0.870	0.467	0.109	0.088
RF	0.702	0.100	1.000	0.854	0.182	0.124	0.488
XGB	0.704	0.250	0.990	0.870	0.385	0.118	0.006
LASSO	0.761	0.300	0.971	0.862	0.414	0.116	0.014
SVM	0.721	0.000	0.990	0.829	0.000	0.126	0.201
KNN	0.577	0.300	0.854	0.764	0.293	0.236	NA

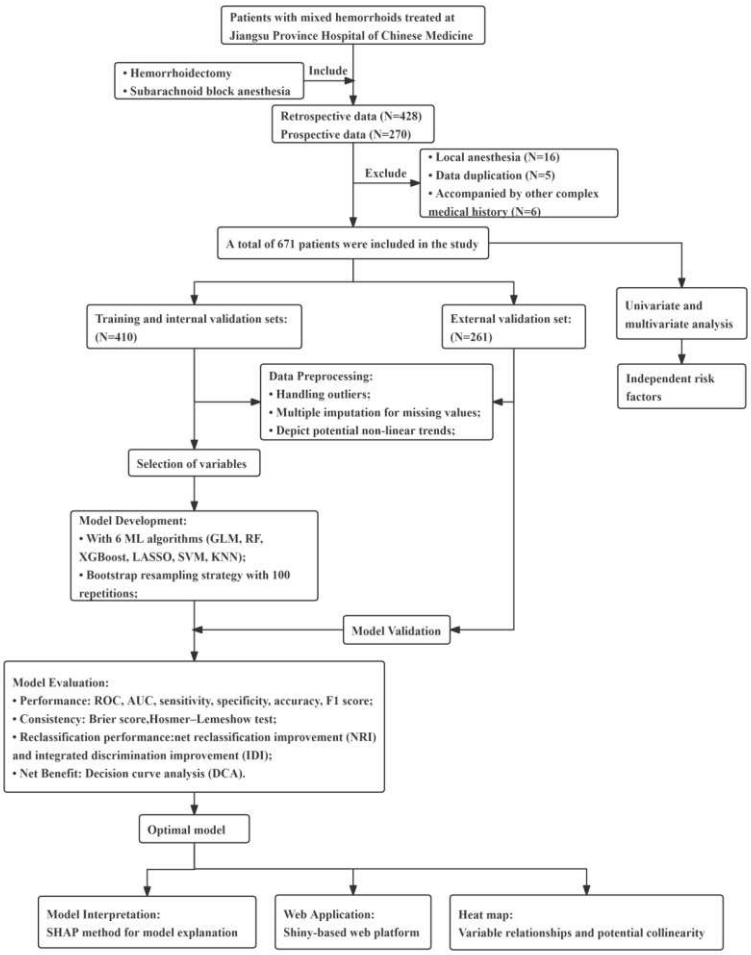
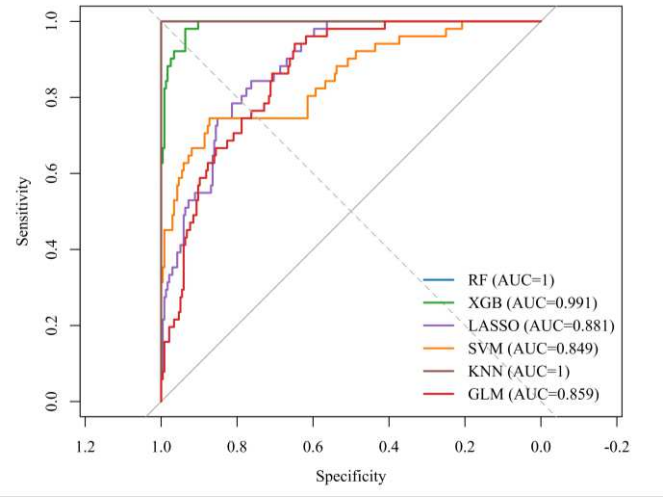


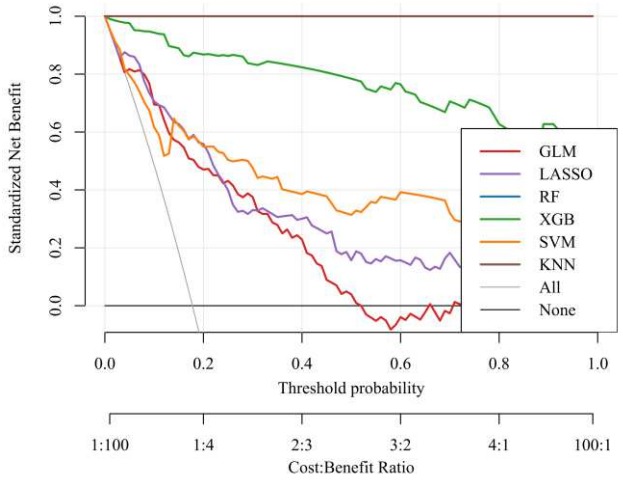
Fig. 1 Flow chart of the study design

GLM: generalized linear model; RF: random forest; XGB:extreme gradient boosting; SVM: support vector machine; KNN: K-nearest neighbors; LASSO: least absolute shrinkage and selection operator; ROC: receiver operating characteristic; AUC: area under curve; SHAP:SHapley additive explanation.



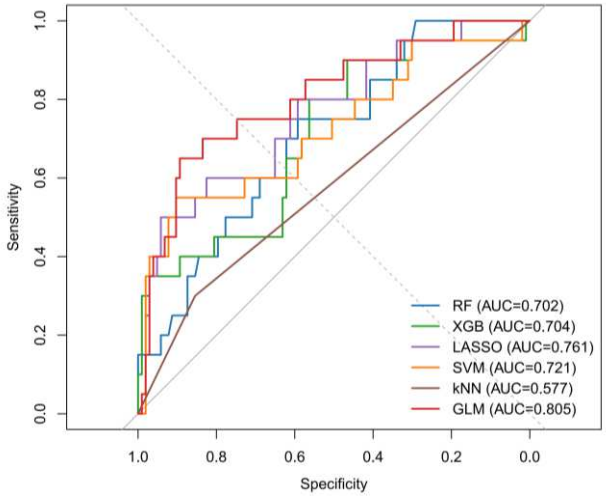
488

489 **Fig. 2 ROC Curves In Training Set**



490

491 **Fig. 3 Training Set DCA**



492

493 **Fig. 4 ROC Curves In Internal Validation Set**

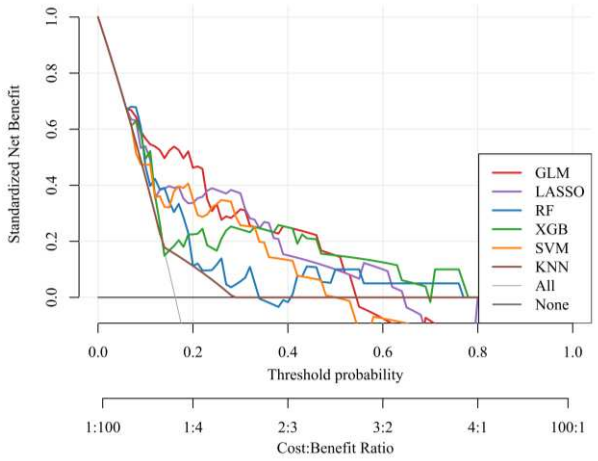


Fig. 5 Internal Validation Set DCA

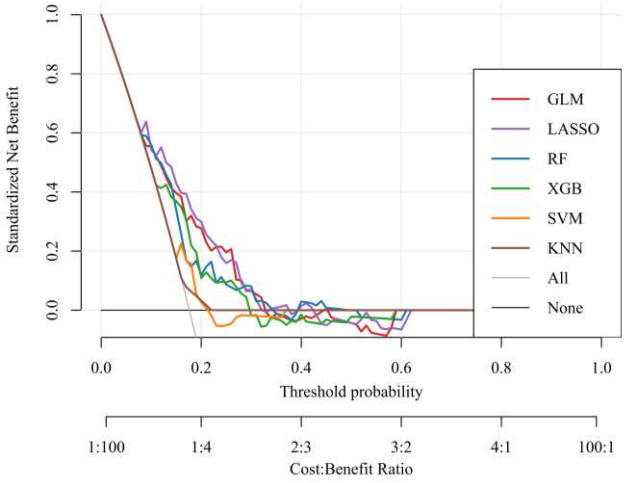


Fig. 6 External Validation Set DCA

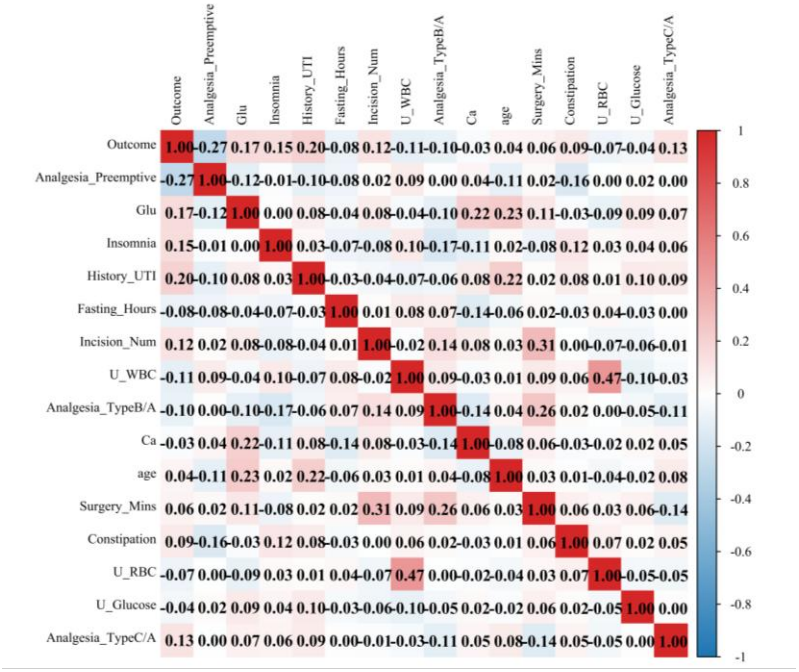


Fig. 7 Correlation Heatmap of the LASSO model

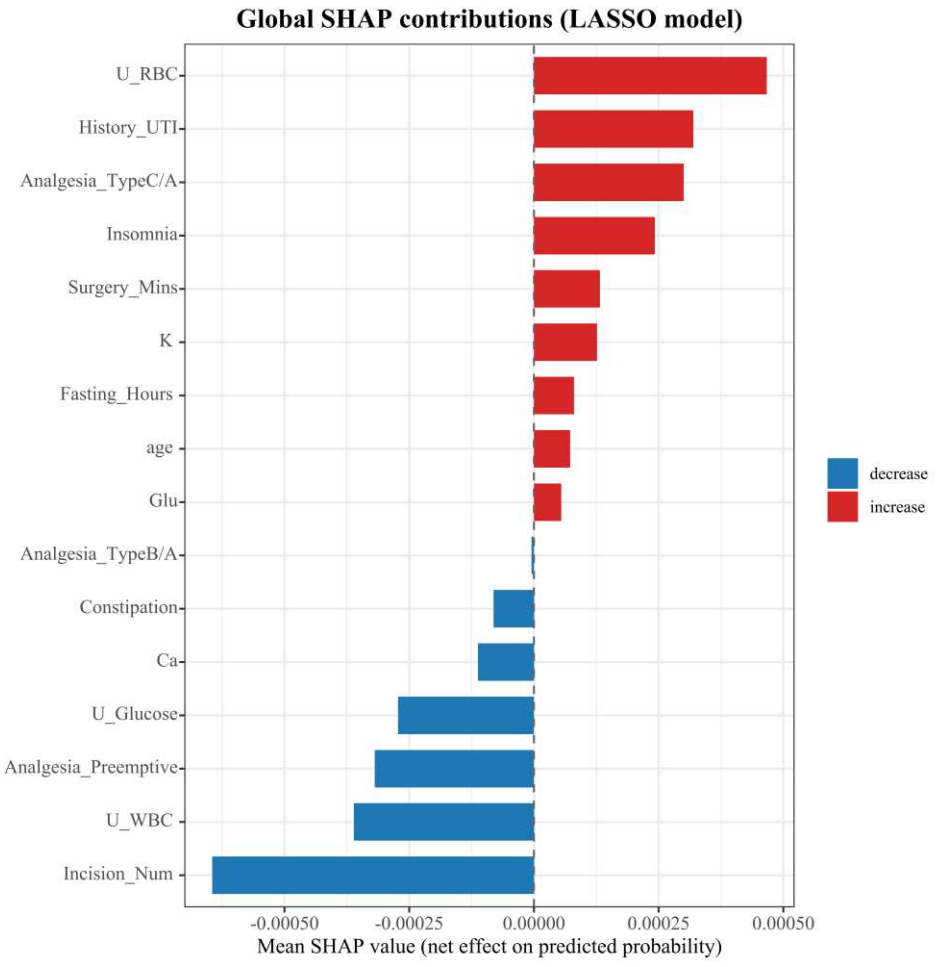


Fig. 8 SHAP waterfall plot (LASSO model)

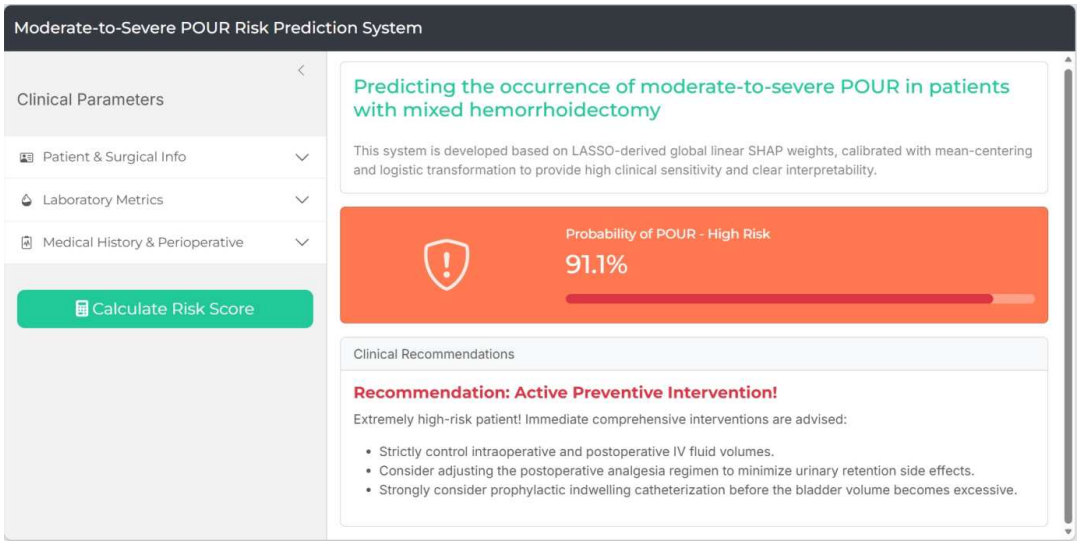


Fig. 9 Web application of predict the risk of moderate-to-severe POUR in individual patient

Ethics Approval and Consent to Participate

The study was approved by the Ethics Committee (Approval No. 2024NL-325-02). Retrospective data were collected from existing medical records, and all data were anonymized, so the requirement for informed consent from patients could be waived. All patients in the prospective data have signed informed consent forms.

Disclosures

The authors declare that they have no relevant financial interests.

Supplementary Materials

All required supplementary materials have been completed and uploaded as separate files.

Data availability

Data will be made available on request.

Funding Statement:

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

[1] Ashburn J. H. (2025). Hemorrhoidal Disease: A Review. *JAMA*, 10.1001/jama.2025.13083. Advance online publication. https://doi.org/10.1001/jama.2025.13083

[2] Toyonaga T, Matsushima M, Sogawa N, Jiang SF, Matsumura N, Shimojima Y et al (2006) Postoperative urinary retention after surgery for benign anorectal disease: potential risk factors and strategy for prevention. Int J Colorectal Dis 21(7):676-682. https://doi.org/10.1007/s00384-005-0077-2

[3] Liu B, Chen Y, Zhang P, Long W, He H, Li X et al (2024) Risk factors for postoperative urinary retention in patients underwent surgery for benign anorectal diseases: a nested case-control study. BMC Anesthesiol 24(1):272. https://doi.org/10.1186/s12871-024-02652-0

[4] Baldini G, Bagry H, Aprikian A, Carli F (2009) Postoperative urinary retention: anesthetic and perioperative considerations. Anesthesiology 110(5):1139-1157. https://doi.org/10.1097/ALN.0b013e31819f7aea

[5] Croghan SM, Mohan HM, Breen KJ, McGovern R, Bennett KE, Boland MR et al (2023) Global Incidence and Risk Factors Associated with Postoperative Urinary Retention Following Elective Inguinal Hernia Repair: The Retention of Urine After Inguinal Hernia Elective Repair (RETAINER I) Study. JAMA Surg 158(8):865-873. https://doi.org/10.1001/jamasurg.2023.2137

[6] Wu JM, Yeh CC, Wei N, Tsai HH, Tseng SM, Chan KC et al (2024) Increased age and the volume of intraoperative fluid administered predict urinary retention after elective inguinal herniorrhaphy. Perioper Med (Lond) 13(1):90. https://doi.org/10.1186/s13741-024-00446-z

[7] Saint S, Trautner BW, Fowler KE, Colozzi J, Ratz D, Lescinskas E et al (2018) A Multicenter Study of Patient-Reported Infectious and Noninfectious Complications Associated with Indwelling Urethral Catheters. JAMA Intern Med 178(8):1078-1085. https://doi.org/10.1001/jamainternmed.2018.2417

[8] Wei JT, Dauw CA, Brodsky CN (2025) Lower Urinary Tract Symptoms in Men: A

Review. JAMA 334(9):809-821. https://doi.org/10.1001/jama.2025.7045

[9] Karavitakis M, Kyriazis I, Omar MI, Gravas S, Cornu JN, Drake MJ et al (2019) Management of Urinary Retention in Patients with Benign Prostatic Obstruction: A Systematic Review and Meta-analysis. Eur Urol 75(5):788-798. https://doi.org/10.1016/j.eururo.2019.01.046

[10] Zhu Y, Wang F, Zhou J, Gu S, Gong L, Lin Y et al (2022) Effect of Acupoint Hot Compress on Postpartum Urinary Retention After Vaginal Delivery: A Randomized Clinical Trial. JAMA Netw Open 5(5):e2213261. https://doi.org/10.1001/jamanetworkopen.2022.13261

[11] Serra-Aracil X, Hidalgo J, Perucho NL, Montesinos CS, Diaz CG, Figueras RV et al (2025) Multicentre, prospective, randomized controlled non-inferiority trial of bladder catheter management in short-duration, minimally invasive colon surgery (The Vesicalcath I-Study). Int J Surg 111(11):8088-8098. https://doi.org/10.1097/JS9.0000000000002962

[12] Huang L, Yin Y, Liao Y, Liu J, Zhu K, Yuan X et al (2022) Risk factors for postoperative urinary retention in patients undergoing colorectal surgery: a systematic review and meta-analysis. Int J Colorectal Dis 37(12):2409-2420. https://doi.org/10.1007/s00384-022-04281-w

[13] Jeong HY, Song SG, Lee JK (2022) Predictors of postoperative urinary retention after semiclosed hemorrhoidectomy. Ann Coloproctol 38(1):53-59. https://doi.org/10.3393/ac.2021.00304.0043

[14] Li Y, Sperrin M, Ashcroft DM, van Staa TP (2020) Consistency of variety of machine learning and statistical models in predicting clinical risks of individual patients: longitudinal cohort study using cardiovascular disease as exemplar. BMJ 371:m3919. https://doi.org/10.1136/bmj.m3919

[15] Pepe M, Longton G, Janes H (2009) Estimation and Comparison of Receiver Operating Characteristic Curves. Stata J 9(1):1.

[16] Kramer AA, Zimmerman JE (2007) Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. Crit Care Med 35(9):2052-2056. https://doi.org/10.1097/01.CCM.0000275267.64078.B0

[17] Kerr KF, Wang Z, Janes H, McClelland RL, Psaty BM, Pepe MS (2014) Net reclassification

indices for evaluating risk prediction instruments: a critical review. Epidemiology 25(1):114-121. https://doi.org/10.1097/EDE.0000000000000018

[18] Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N et al (2010) Assessing the performance of prediction models: a framework for traditional and novel measures.Epidemiology 21(1):128-138. https://doi.org/10.1097/EDE.0b013e3181c30fb2

[19] Porche K, Maciel CB, Lucke-Wold B, Mehkri Y, Murtaza Y, Goutnik M et al (2023) Pre-operative prediction of post-operative urinary retention in lumbar surgery: a prospective validation of machine learning model. Eur Spine J 32(11):3868-3874. https://doi.org/10.1007/s00586-023-07954-4

[20] Shi J, Lin Y, Qin X, Gong Y, Li H (2025) Development and validation of a risk prediction model for postoperative urinary retention after gynecologic abdominal-pelvic surgery. Sci Rep 16(1):2871. https://doi.org/10.1038/s41598-025-32714-3

[21] Jeon C, Hwang S, Cho J (2025) Prediction model for postoperative urinary retention in patients undergoing totally extraperitoneal groin hernia repair. Surgery 185:109549. https://doi.org/10.1016/j.surg.2025.109549

[22] Kim MJ, Lee S, Lee SY, Oh S, Jeon MJ (2024) Development and validation of a prediction model for postoperative urinary retention after prolapse surgery: A retrospective cohort study. BMC Womens Health 24(1):331. https://doi.org/10.1186/s12905-024-03171-3

[23] Duan Y, Du Y, Mu Y, Guan X, He J, Zhang J et al (2025) Development and validation of a novel predictive model for postpancreatectomy hemorrhage using lasso-logistic regression: an international multicenter observational study of 9631 pancreatectomy patients. Int J Surg 111(1):791-806. https://doi.org/10.1097/JS9.0000000000001883

[24] Bunea F, She Y, Ombao H, Gongvatana A, Devlin K, Cohen R (2011) Penalized least squares regression methods and applications to neuroimaging. Neuroimage 55(4):1519-1527. https://doi.org/10.1016/j.neuroimage.2010.12.028

[25] Tibshirani R (1997) The lasso method for variable selection in the Cox model. Stat Med 16(4):385-395.https://doi.org/10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3

[26] Ghuman A, Dawidek MT, Athwal MS, Kasteel N, Brown CJ, Karimuddin AA et al (2022) Prophylactic tamsulosin and urinary retention rates following elective colorectal surgery:

a retrospective cohort study. Int J Colorectal Dis 37(1):209-214. https://doi.org/10.1007/s00384-021-04047-w

[27] Niklasson A, Finan PH, Smith MT, Forsberg A, Dietz N, Kander T et al (2025) The relationship between preoperative sleep disturbance and acute postoperative pain control: A systematic review and meta-analysis. Sleep Med Rev 79:102014. https://doi.org/10.1016/j.smrv.2024.102014

[28] Afolalu EF, Ramlee F, Tang NKY (2018) Effects of sleep changes on pain-related health outcomes in the general population: A systematic review of longitudinal studies with exploratory meta-analysis. Sleep Med Rev 39:82-97. https://doi.org/10.1016/j.smrv.2017.08.001

[29] Odegård SS, Sand T, Engstrøm M, Stovner LJ, Zwart JA, Hagen K (2011) The long-term effect of insomnia on primary headaches: a prospective population-based cohort study (HUNT-2 and HUNT-3). Headache 51(4):570-580. https://doi.org/10.1111/j.1526-4610.2011.01859.x

[30] Buysse DJ (2013) Insomnia. JAMA 309(7):706-716. https://doi.org/10.1001/jama.2013.193

[31] Brownlee M (2001) Biochemistry and molecular cell biology of diabetic complications. Nature 414(6865):813-820. https://doi.org/10.1038/414813a

[32] Vincent AM, Callaghan BC, Smith AL, Feldman EL (2011) Diabetic neuropathy: cellular mechanisms as therapeutic targets. Nat Rev Neurol 7(10):573-583. https://doi.org/10.1038/nrneurol.2011.137

[33] Erasmus, S., Lyu, Z., Zhou, J., Fang, J., & Liang, Y. (2024). Electroacupuncture Mechanisms in Managing Preoperative Anxiety and Postoperative Pain Chronification: A Review. Journal of pain research, 17, 4089–4100. https://doi.org/10.2147/JPR.S498373

[34] The Lancet Respiratory Medicine (2018) Opening the black box of machine learning. Lancet Respir Med 6(11):801. https://doi.org/10.1016/S2213-2600(18)30425-9

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterials.docx