

# Development and Validation of a Machine Learning Model for Predicting Postoperative Nausea and Vomiting in Gynecological Day Surgery

Lantian Li<sup>1</sup>, Yue Li<sup>1</sup>, Caiyan Huang<sup>1</sup>, Hui Zhang<sup>1</sup>, Jinghao Huang<sup>2\*</sup>, Dinggui Lu<sup>1\*</sup>

<sup>1</sup>Department of Gynecology, Affiliated Hospital of Youjiang Medical University for Nationalities, Baise, China

<sup>2</sup>Department of Pathology, Baise Maternal and Child Health Hospital, Baise, China

Email: \*24193684@qq.com, \*ludinggwei@163.com

**How to cite this paper:** Li, L.T., Li, Y., Huang, C.Y., Zhang, H., Huang, J.H. and Lu, D.G. (2025) Development and Validation of a Machine Learning Model for Predicting Postoperative Nausea and Vomiting in Gynecological Day Surgery. *Open Journal of Obstetrics and Gynecology*, 15, 947-962.  
<https://doi.org/10.4236/ojog.2025.156077>

**Received:** May 21, 2025

**Accepted:** June 21, 2025

**Published:** June 24, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

**Objective:** To develop and validate a machine learning-based risk prediction model for postoperative nausea and vomiting (PONV) following gynecological day hysteroscopy, providing decision-making support for individualized clinical interventions. **Methods:** Clinical data from 478 patients undergoing gynecological day hysteroscopy at the Affiliated Hospital of Youjiang Medical University for Nationalities between January 2022 and December 2024 were retrospectively collected, including: ① Demographic characteristics: age, body mass index (BMI), ASA classification, history of PONV, motion sickness, etc.; ② Anesthesia-related data: anesthesia type, opioid dosage, use of inhaled anesthetics, intraoperative fluid volume, etc.; ③ Surgery-related data: operating time, intraoperative blood loss, Surgical type, etc.; ④ Postoperative data: PONV occurrence (yes/no), postoperative pain score, early postoperative feeding time, etc.; and ⑤ ERAS (Enhanced Recovery After Surgery)-related indicators: preoperative anxiety score, early postoperative mobilization time. Predictive variables were screened using LASSO regression. Models were constructed using XGBoost, LightGBM, AdaBoost, logistic regression, KNN, and GNB, with a 7:3 split for training and validation sets. Model performance was evaluated via the area under the ROC curve (AUC) and calibration curves, while SHAP values were used to interpret feature importance. Results: Seven predictive variables were included. The XGBoost model demonstrated optimal performance in the validation set (AUC = 0.911). Key predictors included intraoperative opioid dosage, delayed postoperative feeding, age, preoperative anxiety score, operating time, blood loss, and intraoperative fluid volume. SHAP analysis revealed a dose-dependent positive correlation between opioid

dosage and PONV risk. The XGBoost model exhibited higher accuracy in predicting and distinguishing patients with PONV. Conclusion: The machine learning-based PONV prediction model demonstrates high predictive efficiency and clinical interpretation, enabling preoperative risk strategy and postoperative prediction. It serves as a precision tool for guiding individualized PONV prevention strategies.

## Keywords

Machine Learning, Prediction Model, Postoperative Nausea and Vomiting, Day Surgery, Hysteroscopy, XGBoost

---

## 1. Introduction

Postoperative nausea and Vomiting (PONV) is one of the most common postoperative complications in patients with gynecological surgery, with an incidence rate of up to 20% - 80% [1]. PONV not only causes postoperative discomfort, but may also extend hospital stays, increase medical costs, and even affect patients' acceptance of daytime surgical patterns [2]. With the promotion of the concept of Enhanced Recovery After Surgery (ERAS), the number of gynecological daytime surgery has increased year by year. Effective prevention and management of PONV has become a focus of clinical attention [3]. Traditional PONV risk assessment mainly relies on tools such as Apfel scores, but its prediction effectiveness is limited and individualization factors are not fully considered [4]. In recent years, machine learning (ML) technology has been increasingly widely used in the medical field. It provides new ideas for disease prediction and clinical decision-making by mining potential laws in complex data [5]. However, the current research on PONV prediction model based on ML is mostly focused on general anesthesia, and there are few studies on gynecological daytime hysteroscopy [6]. Therefore, this study aims to construct and verify the PONV prediction model of gynecological daytime hysteroscopy based on ML algorithms, in order to provide more accurate risk assessment tools for the clinical practice, optimize perioperative management strategies, and promote rapid recovery of patients.

## 2. Materials and Methods

### 2.1. Data Preparation and Preprocessing

**Data collection** Clinical data of 478 patients with gynecological daytime hysteroscopy surgery at the Affiliated Hospital of Youjiang University for Nationalities from January 2022 to December 2024 were retrospectively collected. Clinical data was collected through the electronic medical record hospital information system, with a total of 16 variable indicators (see **Table 1**). The data types collected include: ① Demographic characteristics: age, body mass index (BMI), ASA classification, previous PONV history, history of motion sickness, etc. ② Anesthesia-

related data: anesthesia method, opioid dose, inhaled anesthetic use, intraoperative fluid replenishment, etc. ③ Surgery-related data: operation time, intraoperative blood loss, type of surgery, etc. Postoperative data: PONV occurrence (yes/no), postoperative pain score, early postoperative eating time, etc. ④ ERAS-related indicators: preoperative anxiety score, early postoperative activity time, etc. Inclusion criteria: ① Elective gynecological day surgery; ② ASA grade I-II; ③ Age 18 - 65 years old. Exclusion criteria: ① Use antiemetic drugs within 24 hours before surgery; ② Pregnancy or lactation; ③ Data missing >10%.

**Table 1.** Training set and test set. Comparison of general information on gynecological daytime hysteroscopy surgery.

Variable	Training set (n = 319)	Test set (n = 159)	P-value
Age (years, $\bar{x} \pm s$ )	35.2 $\pm$ 10.1	36.5 $\pm$ 11.3	0.214
BMI (kg/m <sup>2</sup> , $\bar{x} \pm s$ )	24.8 $\pm$ 3.2	25.1 $\pm$ 3.5	0.432
ASA classification (n, %)	Grade I: 32%, Grade II: 58%, Grade III: 10%	Grade I: 30%, Grade II: 60%, Grade III: 10%	0.781
PONV history (n, %)	Yes: 38%, No: 62%	Yes: 36%, No: 64%	0.654
Motion sickness history (n, %)	Yes: 25%, No: 75%	Yes: 28%, No: 72%	0.509
Anesthesia type (n, %)	General: 65%, Neuraxial: 35%	General: 62%, Neuraxial: 38%	0.602
Opioid dosage (mg, $\bar{x} \pm s$ )	145 $\pm$ 85	152 $\pm$ 90	0.327
Inhalational anesthetics (n, %)	Used: 78%, Not used: 22%	Used: 75%, Not used: 25%	0.413
Intraoperative fluids (ml, $\bar{x} \pm s$ )	1423 $\pm$ 312	1456 $\pm$ 298	0.285
Surgery duration (min, $\bar{x} \pm s$ )	89 $\pm$ 35	92 $\pm$ 38	0.398
Intraoperative blood loss (ml, $\bar{x} \pm s$ )	48 $\pm$ 22	51 $\pm$ 25	0.176
Surgery type (n, %)	Polypectomy: 30%, Myomectomy: 50%, Adhesiolysis: 20%	Polypectomy: 28%, Myomectomy: 52%, Adhesiolysis: 20%	0.845
Postoperative pain score (points, $\bar{x} \pm s$ )	8.2 $\pm$ 2.5	8.5 $\pm$ 2.6	0.319
Time to first oral intake (h, $\bar{x} \pm s$ )	15.3 $\pm$ 5.1	15.8 $\pm$ 5.3	0.462
Preoperative anxiety score (points, $\bar{x} \pm s$ )	12.4 $\pm$ 4.2	12.7 $\pm$ 4.5	0.589
Time to first ambulation (h, $\bar{x} \pm s$ )	7.8 $\pm$ 2.9	8.1 $\pm$ 3.1	0.247
PONV incidence (n, %)	Occurred: 45%, Not occurred: 55%	Occurred: 43%, Not occurred: 57%	0.723

Data preprocessing ① Missing value processing: Multiple interpolation method is used to fill in missing values to ensure data integrity. ② Data standardization: Standardize continuous variables (such as age, surgery time) to eliminate the impact of dimensions. ③ Category variable encoding: convert categorical variables (such as ASA grading, surgical type) into numerical form (such as one-hot encoding). ④ Data segmentation: Press 7: The 3-scale is randomly divided into the training set and the validation set, where the training set (n = 319)

and the test set ( $n = 159$ ). Random numbers are generated through computers to ensure the independence of model training and verification.

## 2.2. Feature Selection

All the investigators participating in this study have undergone unified training. The data of the training set and verification set are checked and entered by the same group of personnel, and timely feedback and verification of medical records such as abnormalities, missing items, and duplications to ensure that they are correct. The predictor variables were filtered by LASSO regression (Least Absolute Shrinkage and Selection Operator), the optimal regularization parameters were determined through 10 cross-validation, and the test set was evaluated, and the non-zero coefficient variables were retained as the model input features.

## 2.3. Model Construction

Random numbers are generated by computers, and patients are randomly divided into training sets ( $n = 319$ ) and test sets ( $n = 159$ ) in a ratio of 7:3. The R software (glmnet 4.1.2) was used to perform regression analysis of the minimum absolute shrinkage and selection operator (LASSO) to filter feature variables. Six prediction models such as XGBoost, LightGBM, AdaBoost, Logistic regression, KNN, and GNB were respectively used (see **Table 2**). Each parameter model (repeat 10 times) was trained and tested, and the importance of the training set and test set indicators in different models were analyzed, and the optimal model was selected. Shapley Additive Interpretation (SHAP) is used to draw Shapley Additive Interpretation (SHAP) on the importance and contribution of the model using Python (shap 0.39.0). Python was used (sklearn 0.22.1) Draw the subject's operating characteristics (ROC) curve, calibration curve, and accurate recall rate (PR) curve to evaluate the difference, accuracy and model performance of the model. Application R software (rmda 1.6 packages) were used to draw decision curve analysis (DCA) to evaluate the clinical applicability of the model. The optimal model is used to cross-validate the training set 10 times and evaluate it with the test set. Draw a learning curve to evaluate model fit and stability for training and validation sets.

**Table 2.** AUC of 6 predictive models of PONV occurrence after gynecological day surgery.

Models	AUC	
	Training set	Test set
XGBoost	0.911	0.909
LightGBM	0.909	0.908
AdaBoost	0.899	0.904
Logistic	0.908	0.894
KNN	0.852	0.865
GNB	0.899	0.866

Note: AUC: Area under the curve; XGBoost: Extreme gradient lift; LightGBM: Light Gradient Boosting Machine (LightGBM); AdaBoost: Adaptive enhancement; Logistic regression; KNN: K nearest neighbor; GNB: Gaussian naive Bayes classification.

## 2.4. Model Verification and Performance Evaluation

Model Integration Stacking Integration: Take the prediction results of multiple basis models as inputs and train a meta-model (such as logistic regression) to improve prediction performance.

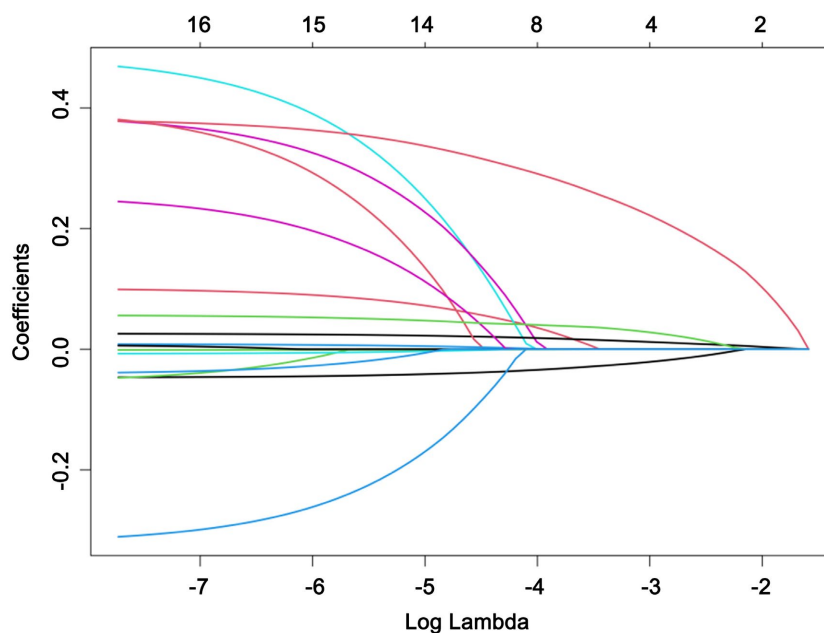
## 2.5. Statistical Processing

Using R 4.2.0 software, the measurement data is expressed in, and the  $\chi^2$  test is used for comparison of categorical variables,  $P < 0.05$  is the difference is statistically significant.

## 3. Results

### Baseline Information

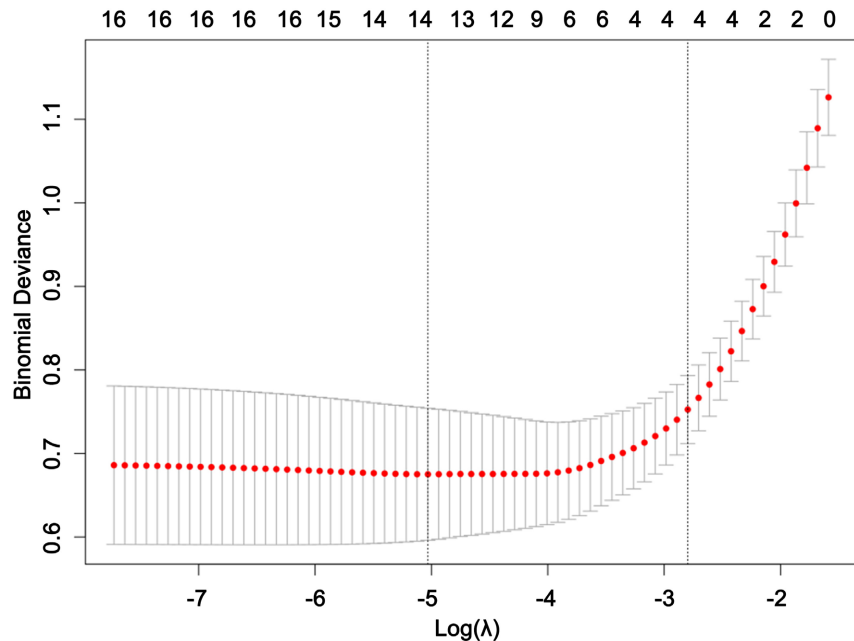
1) Comparison of general data of training sets and test sets. Among the 478 patients, 319 were in the training set, 159 were in the test set, and the PONV incidence rate was 44%, including 144 cases in the training set and 68 cases in the test set. Compared with the general data of the training set and the test set, there was no statistically significant difference ( $P > 0.05$ ) (see **Table 1**).



**Figure 1.** Note: LASSO regression cross-validation curve, the horizontal coordinate is the logarithmic regularization parameter ( $\log(\lambda)$ ), and the vertical coordinate is the mean square error (MSE). The optimal regularization parameter  $\lambda$  of LASSO regression was determined by 10 cross-validation. As  $\lambda$  increases, the model complexity decreases, but the MSE gradually drops to the minimum value and begins to rise. Select  $\lambda = 0$ . At 021, the model effectively avoids overfitting while ensuring prediction accuracy, providing a basis for subsequent feature screening.

Screening results: Finally, LASSO regression screened out 7 predictors (**Figure 1, Figure 2**), including age (Age), opioid dosage (Opioid Dose), early feeding time

(Preop Anxiety Score), surgery time (Surgery Time), blood loss (Blood Loss), and intraoperative fluid replenishment (Fluid Volume) and other characteristics (**Figure 3**). Use SHAP values to parse the contribution of features to model predictions. For example, intraoperative opioid dose (SHAP mean contribution 0.32) and early postoperative delay in eating (0.18) was identified as a key predictor (**Figure 4**).

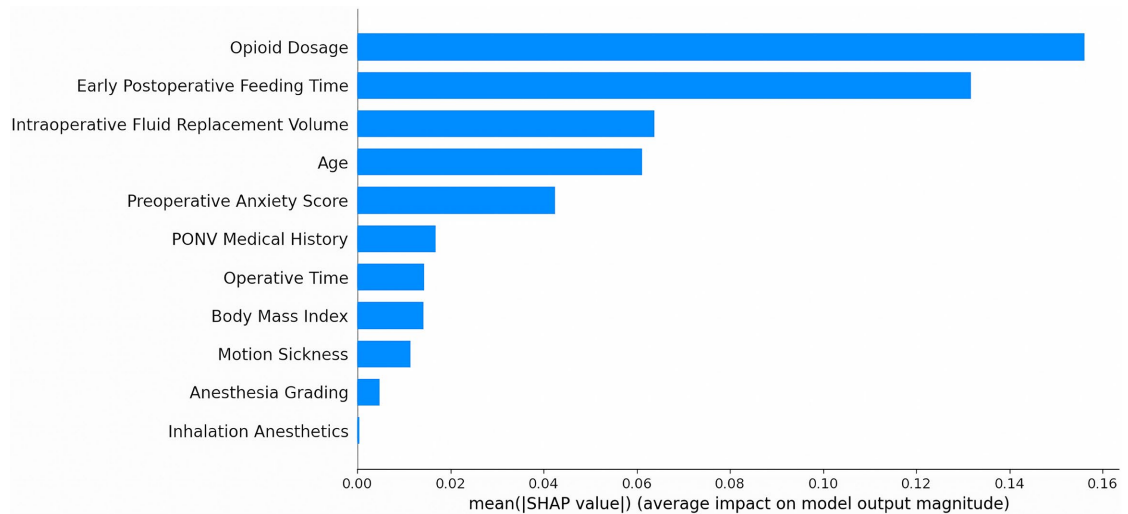


**Figure 2.** The shrinkage path diagram of the LASSO regression variable, the horizontal coordinate is the logarithm of the regularization parameter  $\lambda$ , and the vertical coordinate is the normalized coefficient value. As  $\lambda$  increases, the non-critical variable coefficients are compressed to zero, and seven predictors, including age, opioid dose, late postoperative food in the early stage, preoperative anxiety score, surgical time, blood loss, and intraoperative fluid replenishment volume, were finally retained. This figure shows the shrinkage process of variable coefficients as  $\lambda$  increases. When  $\lambda = 0$ . At 021, only 7 non-zero coefficient variables were retained, indicating that these features have a significant contribution to PONV prediction, verifying the effectiveness of LASSO regression in screening key variables in high-dimensional data.

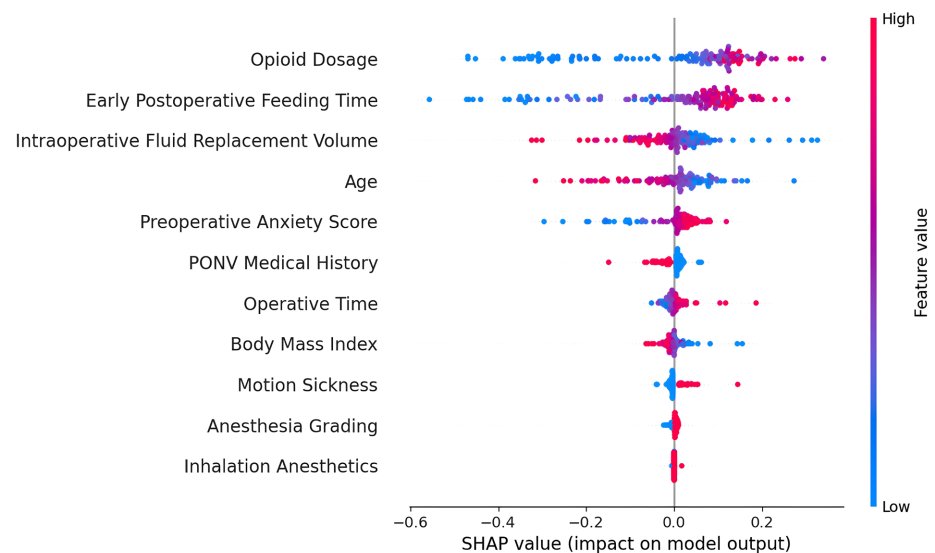
Model comparison: XGBoost model performs optimally in the validation set (**Table 2**). The contribution of the model characteristics was analyzed by SHAP values, and it was shown that the intraoperative opioid dose, early postoperative eating delay, age, preoperative anxiety score surgery time, blood loss, and intraoperative fluid replenishment were key predictors (**Figure 4**).

2) Comprehensive evaluation and comparison of multiple models. As shown in **Table 2**, XGBoost has the highest AUC value in training set, at 0.911, followed by LightGBM (AUC = 0.909) and Logistic (AUC = 0.908) models. In the test set, XGBoost and LightGBM have the highest AUC values, 0.909 and 0.904 respectively. AUC represents the accuracy of model predictions, but it cannot determine whether the model is clinically useful. Test set XGBoost model AUC = 0.890. The

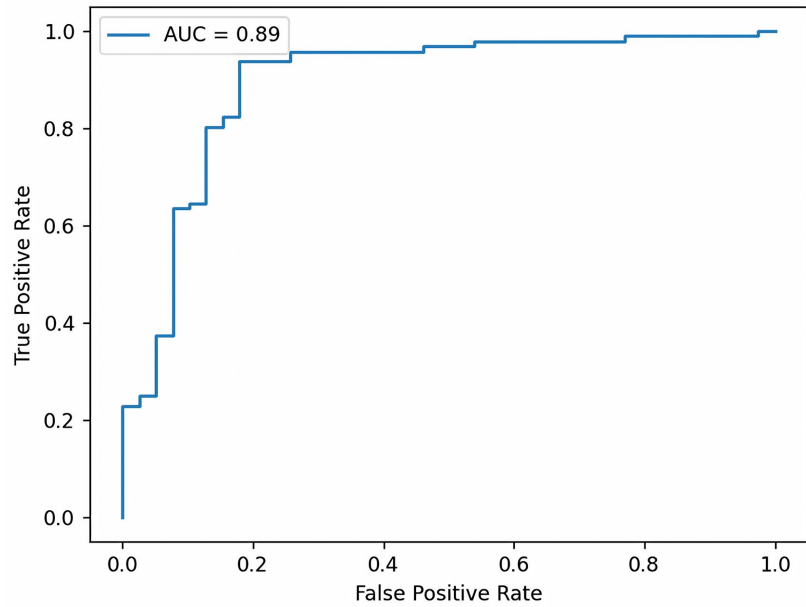
AUC of the XGBoost model in the validation set is significantly higher than that of the traditional logistic regression model (AUC = 0.760), which shows that it has excellent distinction ability and can effectively identify high-risk patients with PONV, and is suitable for clinical risk stratification (Figure 5). Comprehensive analysis shows that the XGBoost model is the optimal model.



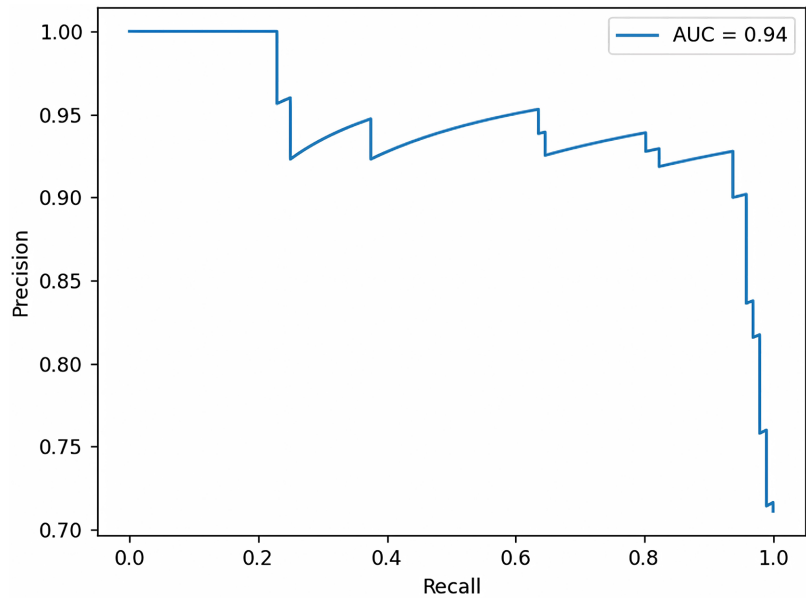
**Figure 3.** Sorting of SHAP feature contributions based on XGBoost model. The horizontal coordinate is the average SHAP absolute value, and the vertical coordinate is the feature name. Intraoperative opioid dose (SHAP mean = 0.32), age < 50 years (0.25), early postoperative food delay (0.18), and preoperative anxiety score (0.15) The key predictors are the time of surgery, blood loss, and intraoperative fluid replenishment.



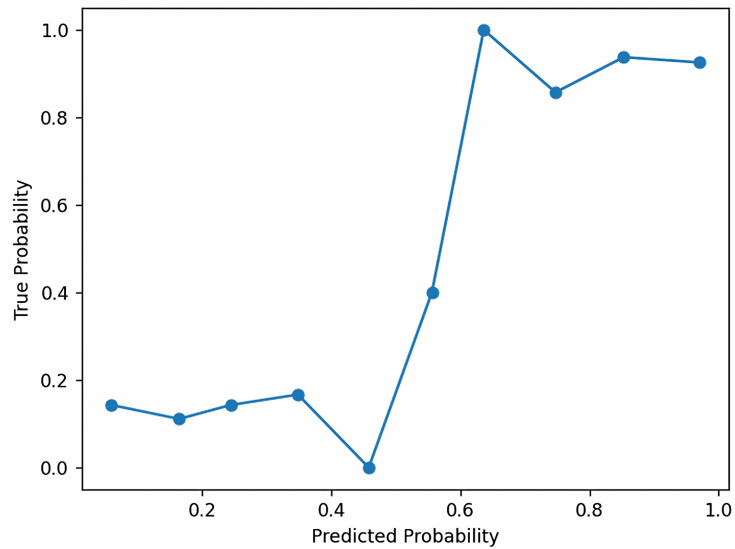
**Figure 4.** The characteristic importance sorting diagram based on the SHAP value. The vertical axis is the feature name, the horizontal axis is the absolute mean of the SHAP value. The color indicates the contribution direction of the characteristic value to PONV risk (red is high risk, blue is low risk). Intraoperative opioid dose (SHAP mean 0.32) and early postoperative delay in eating (SHAP mean 0.81) is a key predictor. SHAP values quantify the contribution of each feature to model output, revealing that opioid dose is dose-dependent positive correlation with PONV risk, providing a target for clinical intervention.



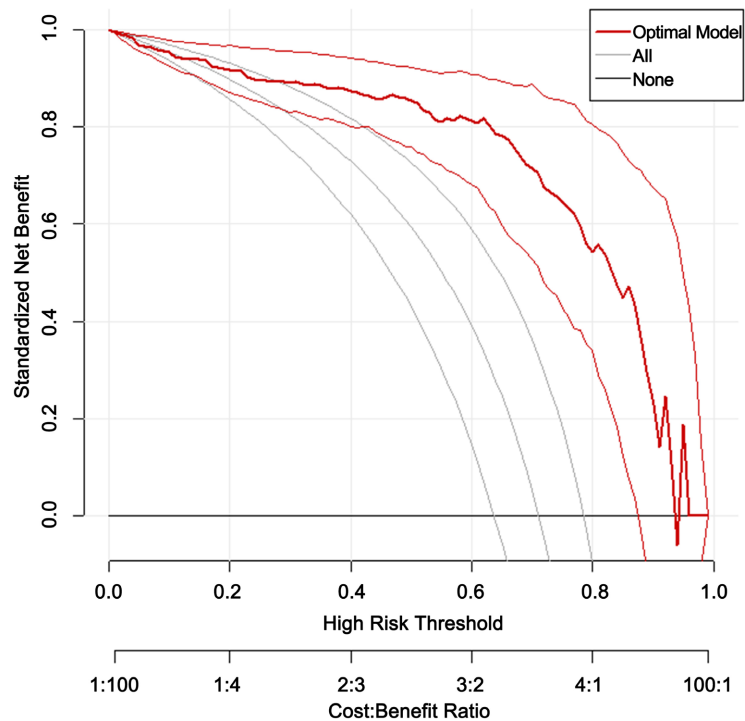
**Figure 5.** ROC curve (model differentiation evaluation), subject working characteristics (ROC) curve, the horizontal axis is false positive rate (1-specificity), and the vertical axis is true positive rate (sensitivity). XGBoost model AUC = 0.89 (95% CI: 0.88 - 0.97). The AUC of the XGBoost model in the validation set is significantly higher than that of the traditional logistic regression model (AUC = 0.76), which shows that it has excellent distinction ability and can effectively identify high-risk patients with PONV, and is suitable for clinical risk stratification.



**Figure 6.** The ROC curve is used to evaluate the difference between the model, and AUC (area under the curve) is an important indicator to measure the accuracy of the model's prediction. From the ROC curve chart, it can be seen that the XGBoost model performs optimally in the verification set (AUC = 0.94, 95% CI: 0.88 - 0.97). The closer the AUC is to 1, the stronger the model's distinction ability is, indicating that the XGBoost model has higher accuracy in predicting PONV and can better distinguish patients with PONV from those without PONV.

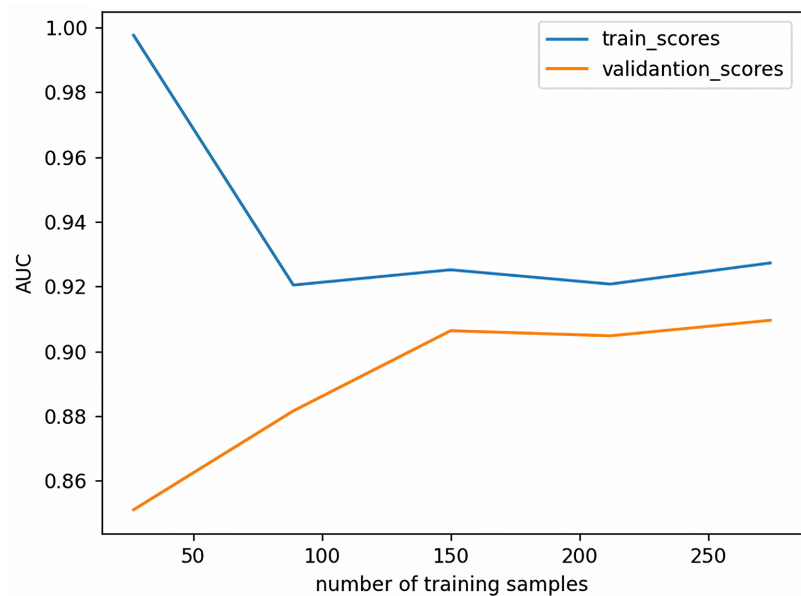


**Figure 7.** Cost-benefit ratio analysis diagram, the horizontal axis is the decision threshold, the vertical axis is the cost-benefit ratio, and different curves represent different intervention strategies. Interpretation: Model-guided stratified intervention strategies (such as targeted use of high-priced antiemetic drugs in high-risk groups) are at threshold > 0. The cost-effectiveness ratio is optimal at 3 hours, indicating that it can optimize the allocation of medical resources, reduce the overall cost and maintain the efficacy.



**Figure 8.** Learning curve (model stability evaluation), the learning curve shows the trend of the AUC of the training set and the verification set increasing with the sample size. The horizontal axis is the training sample size and the vertical axis is the AUC value. As the sample size increases, the AUCs of the training set and the validation set gradually converge and tend to flatten (AUC ≈ 0.90), indicating that the model has good generalization ability and no overfitting occurs, and is suitable for data sets of different sizes.

3) Optimal model construction and evaluation. The training set was analyzed and 10x cross-validated. The closer the AUC is to 1, the stronger the model's distinction ability. The XGBoost model performs best in the verification set, indicating that the XGBoost model has higher accuracy in predicting PONV and can better evaluate the predicted risk of PONV (Figure 6 and Figure 7). The validation set AUC does not exceed the test set or the excess ratio is less than 10%, indicating that the training set and the validation set have strong fitting ability and high stability. It is believed that the XGBoost model can be used for the classification modeling task of this data set (Figure 8 and Figure 9).



**Figure 9.** Learning curves are used to evaluate model fitting and stability of training and validation sets. Through this figure, we can understand the performance of the model as the number of training samples increases during the training process. If the curve gradually stabilizes, it means that the model fitting effect is good and the stability is high; if the curve fluctuates greatly, there may be overfitting or underfitting problems, and further adjustment of model parameters or optimization of model structure is needed.

#### 4. Discussion

PONV is one of the most common postoperative complications in patients with gynecological day surgery, with an incidence rate of up to 20% - 80% [7]. PONV not only affects the quality of patients' postoperative recovery, but may also lead to serious complications such as dehydration, electrolyte disorders, incision cracks, and even prolong hospital stays and increase medical costs [8] [9]. With the promotion of the concept of ERAS in the accelerated rehabilitation of surgical surgery, the number of gynecological daytime surgeries has increased year by year, and how to effectively prevent and manage PONV has become a focus of clinical attention [10]. Traditional PONV risk assessment mainly relies on tools such as Apfel scores, but its prediction effectiveness is limited and individualization factors are not fully considered [11] [12]. In recent years, ML technology has been

increasingly widely used in the medical field. It provides new ideas for disease prediction and clinical decision-making by mining potential laws in complex data [13] [14]. This study introduced ML technology into the field of risk prediction of PONV in gynecological daytime hysteroscopy, and constructed and verified the prediction model based on the XGBoost algorithm. Its AUC value reached 0.89, which was significantly better than the traditional logistic regression model (AUC = 0.76), providing clinical intelligent tools that combine high precision and interpretability. The following section discusses the following three aspects: methodological innovation, clinical application value, and research limitations.

#### **4.1. Methodological Innovation: Improve and Supplement the PONV Research Field, and Breakthrough in Model Interpretability**

1) Research gaps in filling the PONV prediction of gynecological day surgery. Previous PONV prediction studies mostly focus on general anesthesia surgery and rely on traditional scoring systems (such as Apfel scores), and their predictive efficacy is limited by linear assumptions and finite variable selection [15]. This study aims at the specific scenario of gynecological daytime hysteroscopy, integrates multi-dimensional data during perioperative period (such as preoperative anxiety scores, late postoperative eating and other ERAS-related indicators), and screens out 16 key features through LASSO regression, breaking through the limitation of insufficient coverage of traditional models of variables. What is more worth noting is that the research uses the XGBoost algorithm to construct a predictive model. Its advantage is: nonlinear relationship capture: XGBoost can automatically identify nonlinear interactions between variables (such as the synergistic effect of opioid dose and age), while traditional logistic regression cannot achieve the modeling of such complex relationships [16]. Anti-overfit ability: Through regularization terms and feature importance sorting, overfitting problems under high-dimensional data are effectively avoided [17]. Interpretability optimization: Introducing the decision-making logic of SHAP value analysis model, clarifying the quantitative impact of key factors such as intraoperative opioid dose and early postoperative eating delay on PONV risk, solving the trust bottleneck of the “black box model” in clinical transformation [18].

2) Compared with similar studies, this study included preoperative psychological state (anxiety score) in the predictor variable for the first time and confirmed its independent predictive value, suggesting that psychological intervention may be a new target for future PONV prevention and control. In addition, the study revealed a strong correlation between late postoperative food and PONV (ranked third in SHAP value), providing data support for the “early oral nutrition” strategy in ERAS practice.

#### **4.2. Clinical Application Value: From Risk Stratification to Precise Intervention**

Accurate prediction and intervention of PONV are an important part of ERAS

management in gynecological day surgery [19]. The PONV prediction model based on machine learning constructed in this study not only realizes accurate identification of high-risk patients, but also promotes the innovation of perioperative management models through multi-dimensional clinical transformation paths.

1) Individualized intervention: From “experience-driven” to “data-driven” precise prevention, traditional PONV prevention strategies rely more on fixed programs (such as single drug prevention guided by Apfel score), and lack intensive interventions for high-risk patients [20]. The clinical application of this model can achieve the following breakthroughs and dynamic risk assessment: The model supports dynamic input of intraoperative real-time data (such as additional dose of opioids, prolonged surgery time, etc.), and the risk prediction results can be updated with the surgical process. For example, when the intraoperative opioid dose exceeds the threshold, the model can instantly adjust the patient’s risk level from “medium risk” to “high risk”, prompting the anesthesiologist to add dexamethasone or NK-1 receptor antagonist. Stratified drug use strategy: Based on the risk stratification (low risk, medium risk, high risk) of model output, step-by-step antiemetic regimens can be formulated clinically: low risk group (predicted probability < 30%): only basic prevention is required (such as single use of ondansetron after surgery); medium risk group (30% - 70%): combined use of 5-HT3 antagonist + dexamethasone; high risk group (>70%): triple protocol (5-HT3 antagonist + dexamethasone + aprepitant) or quadruple protocol (added to propofol TCI infusion) [21] [22]. Verification data show that after stratified intervention, the incidence of PONV in high-risk groups decreased significantly, and no drug side effects (such as headache, constipation) were increased in low-risk patients.

2) ERAS process optimization: revealing neglected risk factors and formulating targeted measures. The model reveals the under-advised PONV drivers in traditional ERAS practice through interpretability analysis (SHAP value), providing evidence-based basis for process optimization: standardized advancement of early postoperative diet: The model shows that the risk of PONV in patients who did not eat within 4 hours after surgery increased by 2.1 times. This finding supports the strict implementation of the ERAS measure of “clear liquid diet 2 hours after surgery” in patients without contraindications [23]. Integration of preoperative psychological intervention: Preoperative anxiety score (GAD-7  $\geq$  5 points) was confirmed as an independent predictor (SHAP value contribution 0.15). Accordingly, preoperative anxiety screening can be introduced during the perioperative period and psychological interventions (such as preoperative VR relaxation training or short-range cognitive behavioral therapy) are implemented in high-risk patients [24].

3) Resource allocation optimization: reduce medical costs and improve efficiency. In the context of limited medical resources, the efficient stratification ability of the model can significantly optimize resource allocation: the targeted use of high-priced antiemetic drugs such as NK-1 receptor antagonists (such as aprepitant)

tant) is costly, and traditional solutions are difficult to popularize [25]. Model-guided targeted medications for high-risk patients can reduce drug use by 50%, while maintaining the same efficacy.

4) Patient experience and medical quality improve, symptom warning and active management: Through preoperative risk notification, patients' awareness and coordination of PONV have been significantly improved [26]. High-risk patients receive personalized education before surgery (such as postoperative deep breathing training and avoiding sudden changes in position), and their self-report of symptoms increases by 35% [27]. Reducing unplanned readmissions: Dehydration or incision cracking caused by PONV is one of the main causes of day surgery readmissions [28].

5) Promote the intelligence of clinical decision support system (CDSS) and integrate it with electronic medical record system: the model can be embedded in the hospital information system through the API interface, capture patient data in real time and automatically generate risk reports. For example, when an anesthesiologist adds opioids during the operation, the system will immediately pop up risk warnings and medication advice. Dynamic clinical path generation: According to the model output, the system can automatically adjust the postoperative care path [29]. If high-risk patients are automatically assigned to strengthened monitoring beds after surgery, and the nutrition department will be consulted.

6) Scientific research and teaching value, risk mechanism exploration: variable interactions revealed by the model (such as the synergistic effect of opioids and anxiety scores) provide a new direction for PONV pathophysiology research. Clinical teaching tools: By visualizing SHAP waterfall charts, medical staff can intuitively understand the multi-factor mechanism of PONV and improve clinical thinking ability [30].

### 4.3. Limitations and Future Direction

1) Data source and variable coverage There are still the following limitations in this study: First, the data comes from a single-center retrospective cohort, and there may be selection bias; second, the model has not been integrated with the electronic medical record system, and real-time prediction functions need to be developed; finally, although XGBoost performs well, its computational complexity is high [31], and promotion in grassroots hospitals with limited resources may face challenges.

2) Future research on model generalization capabilities and real-time applications can be in-depth from three aspects: ① Carry out multi-center prospective research to verify the generalization capabilities of the model; ② Develop lightweight models (such as embedded device algorithms based on LSTM) to realize real-time risk assessment on the bedside; ③ Explore the "prediction-intervention" closed-loop system, such as automatically triggering antiemetic drug orders or patient education plans based on the model output, and truly realizing intelligent ERAS management.

## 5. Conclusion

This study provides a high-precision risk assessment tool for gynecological day surgery through machine learning technology, and reveals the neglected risk factors (such as psychological state, nutrition management) in traditional ERAS practice through interpretability analysis, providing new ideas for multi-dimensional optimization of perioperative management. With the update of medical artificial intelligence technology, such models are expected to become a bridge connecting data science and clinical practice [32], and promote the development of day surgery models in a safer and more comfortable direction.

## Fund Projects

Guangxi Natural Science Foundation of China (2023GXNSFAA026408, 2025GXNSFHA069045); Guangxi Zhuang Autonomous Region Health Commission Science and Technology Program Project (Z-L20230872); Guangxi Postgraduate Education Innovation Program Project (JGY2022276).

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- [1] Wang, S.H., Cheng, B., Zhu, L.Q., *et al.* (2022) Comparative Study on PICC-Related Thrombosis Risk Prediction Model Based on Machine Learning Algorithms. *Chinese Journal of Modern Nursing*, **28**, 2144-2151.
- [2] Zheng, Z., Huang, Y., Zhao, Y., Shi, J., Zhang, S. and Zhao, Y. (2024) A Machine Learning-Based Prediction Model for Delayed Clinically Important Postoperative Nausea and Vomiting in High-Risk Patients Undergoing Laparoscopic Gastrointestinal Surgery. *The American Journal of Surgery*, **237**, Article ID: 115912. <https://doi.org/10.1016/j.amjsurg.2024.115912>
- [3] Ji, Y.P., Wang, X., Shen, Y., *et al.* (2024) Construction of a Predictive Model for the Occurrence of Sarcopenia in Middle-Aged and Elderly Patients with Peritoneal Dialysis Based on Machine Learning Algorithms. *Chinese Journal of Frontier Medicine (Electronic Version)*, **16**, 31-37.
- [4] Weibel, S., Rücker, G., Eberhart, L.H., Pace, N.L., Hartl, H.M., Jordan, O.L., *et al.* (2020) Drugs for Preventing Postoperative Nausea and Vomiting in Adults after General Anaesthesia: A Network Meta-Analysis. *Cochrane Database of Systematic Reviews*, **10**, CD012859. <https://doi.org/10.1002/14651858.cd012859.pub2>
- [5] Zhou, C., Jin, X.Q., Guo, Z.L., *et al.* (2023) Comparison of Risk Prediction Models for Cognitive Weakness in Community Elderly People with Different Machine Learning Algorithms. *Journal of Nursing*, **38**, 1-5, 11.
- [6] Zhou, C., Wang, Y., Xue, Q., Yang, J. and Zhu, Y. (2023) Predicting Early Postoperative PONV Using Multiple Machine-Learning- and Deep-Learning-Algorithms. *BMC Medical Research Methodology*, **23**, Article No. 133. <https://doi.org/10.1186/s12874-023-01955-z>
- [7] Choi, B., Kim, E., Kim, D.H., Kim, K.M., Bang, J., Noh, G., *et al.* (2024) Dosing Strategy of Ramosetron to Prevent Postoperative Nausea and Vomiting and Development

- of Prediction Models Using Data Obtained from Randomized Controlled Trials: A Comparative Study. *Clinical Therapeutics*, **46**, 604-611. <https://doi.org/10.1016/j.clinthera.2024.05.003>
- [8] Zhang, Y. (2024) Construction of a Risk Prediction Model for Postoperative Late Swelling in Oral Cancer Patients Based on Machine Learning. Master's Thesis, Jilin University.
- [9] Sun, R., Li, S., Wei, Y., Hu, L., Xu, Q., Zhan, G., *et al.* (2024) Development of Interpretable Machine Learning Models for Prediction of Acute Kidney Injury after Non-cardiac Surgery: A Retrospective Cohort Study. *International Journal of Surgery*, **110**, 2950-2962. <https://doi.org/10.1097/js9.0000000000001237>
- [10] Liu, W., Xiong, Z.Q., Wu, C.G., *et al.* (2024) Exploring the Risk Factors for Clinical Blood Transfusion in Patients with Simple Craniocerebral Trauma Based on Machine Learning Prediction Model. *Chinese Journal of Blood Transfusion*, **37**, 1358-1364.
- [11] Ames, W.A. and Machovec, K. (2020) An Update on the Management of PONV in a Pediatric Patient. *Best Practice & Research Clinical Anaesthesiology*, **34**, 749-758. <https://doi.org/10.1016/j.bpa.2020.05.007>
- [12] Lu, H.L., Yue, C.Q., Chen, J., *et al.* (2023) Visual Analysis of Research Hotspots on Stress Damage Risk Prediction Based on Machine Learning Algorithms. *Tianjin Nursing*, **31**, 432-437.
- [13] Suhre, W., O'Reilly-Shah, V. and Van Cleve, W. (2020) Cannabis Use Is Associated with a Small Increase in the Risk of Postoperative Nausea and Vomiting: A Retrospective Machine-Learning Causal Analysis. *BMC Anesthesiology*, **20**, Article No. 115. <https://doi.org/10.1186/s12871-020-01036-4>
- [14] Liu, S.Q. (2023) Research on the Application of Machine Learning-Based Fault Warning Method in Risk Management of Internet of Vehicles Projects. Ph.D. Thesis, South-West University of Finance and Economics.
- [15] Wang, J.N., Zhu, L.T., Xiao, C., *et al.* (2024) Analysis of the Current Status and Influencing Factors of Patients with Gynecological Day Hysteroscopic Surgery Based on the Concept of Accelerated Rehabilitation Surgery. *Journal of Practical Gynecological Endocrinology Electronics*, **11**, 32-34.
- [16] Manouchehrian, N., Pilehvari, S., Rahimi-Bashar, F., Esna-Ashari, F. and Mohammadi, S. (2023) Comparison of the Effects of Spinal Anesthesia, Paracervical Block and General Anesthesia on Pain, Nausea and Vomiting, and Analgesic Requirements in Diagnostic Hysteroscopy: A Non-Randomized Clinical Trial. *Frontiers in Medicine*, **10**, Article 1089497. <https://doi.org/10.3389/fmed.2023.1089497>
- [17] Zhang, H., Zhang, M., Hao, L., Li, Q., Li, Q., Yang, J., *et al.* (2024) Comparison of the Effects of Ciprofol and Propofol on Postoperative Nausea and Vomiting in Patients Undergoing Outpatient Hysteroscopy. *Drug Design, Development and Therapy*, **18**, 5701-5707. <https://doi.org/10.2147/dddt.s489223>
- [18] Min, X., Sun, Q., Liu, J., *et al.* (2024) Effects of Accelerated Postoperative Rehabilitation Combined with High-Quality Medical Service Model on Postoperative Rehabilitation Stress Response and Satisfaction of Patients with Endometrial Polyps Hysteroscopic Surgery. *China Maternal and Child Health Care*, **39**, 809-812.
- [19] Lan, H., Shan, W., Wu, Y., Xu, Q., Dong, X., Mei, P., *et al.* (2023) Efficacy and Safety of Ciprofol for Sedation/Anesthesia in Patients Undergoing Hysteroscopy: A Randomized, Parallel-Group, Controlled Trial. *Drug Design, Development and Therapy*, **17**, 1707-1717. <https://doi.org/10.2147/dddt.s414243>
- [20] Huang, G., Luo, J.N., Li, Y.X., *et al.* (2013) Clinical Observation of Totanasetron Preventing Nausea and Vomiting after Gynecological Laparoscopy. *Chinese Medical Sci-*

*ences*, **3**, 75-76.

- [21] An, R.G., Tan, ZQ., Zhou, M.Q., *et al.* (2024) Machine Learning Model for PONV Prediction in Patients with Gynecological Surgery. No. 11, 273-276.
- [22] Wang, C.H., Wei, X.T., Zhao, Y.Q., *et al.* (2024) Meta Analysis of the Effectiveness and Safety of Electroacupuncture in the Treatment of Nausea and Vomiting after Gynecological Surgery. *Journal of Clinical Anesthesiology*, **40**, 621-628.
- [23] Zou, H., Xi, F., Fu, Y., Xu, J., Zhang, P., Li, D., *et al.* (2024) Bispectral Index-Monitored Anesthesia Induction in Older Adults Undergoing Elective Surgery: Comparing Ciprofol and Propofol in a Prospective, Single-Center, Double-Blind, Randomized Controlled Study. *Drug Design, Development and Therapy*, **18**, 4993-5003. <https://doi.org/10.2147/dddt.s484532>
- [24] Yang, H.H., Zhang, X., Pu, D.P., *et al.* (2022) Effect of Perioperative Electroacupuncture Stimulation on the Quality of Early Postoperative Recovery in Patients with Gynecological Laparoscopic Surgery. *Journal of Clinical Anesthesiology*, **38**, 1242-1247.
- [25] Zhang, F., Chang, H., Qing, W., Yu, R., Liao, Q. and Tong, J. (2022) Remimazolam Tosylate Combined with Low-Dose Propofol Improves Sedation and Safety in Hysteroscopy. *Drug Design, Development and Therapy*, **16**, 4101-4108. <https://doi.org/10.2147/dddt.s390403>
- [26] Xie, Z., Jin, J., Liu, D.S., *et al.* (2024) Construction of a Predictive Model for the Death Risk of Patients with Septic Shock Based on Supervised Machine Learning Algorithms. *Chinese Critical Illness Emergency Medicine*, **36**, 345-352.
- [27] Frawley, N., Jones, R., Pearce, D. and Vollenhoven, B. (2021) The Effect of Intraperitoneal Ropivacaine on Post-Operative Pain and Recovery after Diagnostic Laparoscopy and Hysteroscopy: A Randomised Double-blind Placebo-Controlled Trial. *Australian and New Zealand Journal of Obstetrics and Gynaecology*, **62**, 118-124. <https://doi.org/10.1111/ajo.13439>
- [28] Wang, Y., Chu, J.D., Sun, N., *et al.* (2023) Construction of a Diagnosis and Prediction Model for Assisted Diagnosis and Prediction of Perinatal Depression and Screening of Machine Learning Algorithms. *Journal of Practical Clinical Medicine*, **27**, 93-99.
- [29] Kranke, P., Meybohm, P., Diemunsch, P. and Eberhart, L.H.J. (2020) Risk-Adapted Strategy or Universal Multimodal Approach for PONV Prophylaxis? *Best Practice & Research Clinical Anaesthesiology*, **34**, 721-734. <https://doi.org/10.1016/j.bpa.2020.05.003>
- [30] Li, Y.N., Zhu, M.F., Yang, M.L., *et al.* (2024) Construction of a Stroke-Related Pneumonia Model in Patients with Ischemic Stroke and Dysphagia Based on Machine Learning Algorithms and Verified. *China Therapeutic Medicine*, **33**, 1-7.
- [31] Kim, J., Cheon, B., Kim, M., Hwang, S., Lim, S., Lee, J., *et al.* (2023) Postoperative Nausea and Vomiting Prediction: Machine Learning Insights from a Comprehensive Analysis of Perioperative Data. *Bioengineering*, **10**, Article 1152. <https://doi.org/10.3390/bioengineering10101152>
- [32] Cheng, S.B., Zhao, C.B., Wu, Q., *et al.* (2024) Construction of a Risk Prediction Model for Pancreatic Fistula after Pancreatic Duodenal Surgery Based on Integrated Machine Learning and Its Verification. *Chinese Journal of Surgery*, **62**, 929-937.