International Journal of General Medicine

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

# Temporal Dynamic Machine Learning Prediction of Postoperative Gastrointestinal Dysfunction Duration in Esophageal Cancer: Integrating Preoperative and Perioperative Routine Blood Temporal Data and Dynamic Ultrasonic Features

Lei Jin[1], Yanxiang Chang[1], Hao Liu[1], Jie Guo[2]

[1]Department of Cardiothoracic Surgery, The First Affiliated Hospital of Xi'an Medical University, Xi'an, 710077, People's Republic of China; [2]Department of Gastroenterology, The First Affiliated Hospital of Xi'an Medical University, Xi'an, 710077, People's Republic of China

Correspondence: Jie Guo, Email guojie9222023@163.com



**Objective:** Gastrointestinal dysfunction following esophageal cancer surgery represents a prevalent postoperative complication. This study aims to develop a time-dynamic machine learning model to predict the duration of postoperative gastrointestinal dysfunction (POGID) by integrating preoperative and perioperative continuous blood data with dynamic ultrasound characteristics, thereby facilitating early clinical intervention.

**Methods:** A retrospective cohort of 826 patients who underwent radical esophagectomy between 2017 and 2024 was enrolled and stratified into a training set (70%), a validation set (15%), and a test set (15%). Predictive variables encompassed baseline demographic and clinical data, blood routine parameters at five distinct time points, and ultrasound features at three time points. Four machine learning models were constructed: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Temporal Convolutional Network (TCN), and Random Forest (RF). Model performance was evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Coefficient of Determination ($R^2$), and Mean Absolute Percentage Error (MAPE). Feature importance was assessed via SHapley Additive exPlanations (SHAP) analysis.

**Results:** The LSTM model demonstrated superior predictive performance on the test set, achieving an MAE of $1.23 \pm 0.31$ days, an RMSE of $1.56 \pm 0.42$ days, an $R^2$ of $0.78 \pm 0.06$, and an MAPE of $12.3\% \pm 3.1\%$, significantly outperforming the RF model (all $P < 0.001$). The top five influential predictors were postoperative day 1 white blood cell count, preoperative day 1 antral cross-sectional area, postoperative day 3 platelet count, Tumor Node Metastasis (TNM) stage, and postoperative day 2 intestinal peristalsis frequency. Subgroup analyses confirmed the model's robust predictive capability, with $R^2$ values ranging from 0.72 to 0.83.

**Conclusion:** The time-dynamic LSTM model, which integrates continuous blood data, ultrasound features, and baseline characteristics, accurately predicts POGID duration and identifies actionable intervention targets. This model can be integrated into clinical decision support systems to optimize perioperative management and enhance postoperative recovery.

**Keywords:** esophageal cancer, postoperative gastrointestinal dysfunction, temporal data, ultrasonic dynamic features, machine learning, prediction model

## Introduction

Cancer is one of the most prevalent malignant tumors of the digestive system globally, characterized by high incidence and mortality rates.[1,2] Radical surgical resection remains the primary treatment for localized esophageal cancer; however, it involves significant surgical trauma and gastrointestinal reconstruction, often resulting in a range of postoperative complications.[3] Postoperative gastrointestinal dysfunction (POGID) stands as one of the most common complications, primarily manifesting as delayed gastric emptying, intestinal paralysis, abdominal distension, and loss of appetite.[4]

https://doi.org/10.2147/IJGM.S588931
Received: 3 January 2026
Accepted: 11 April 2026
Published: 23 April 2026

According to reports, the incidence of POGID among esophageal cancer patients ranges from 35% to 60%, with a duration of 3 to 14 days.[4–6] Prolonged POGID not only elevates the risk of anastomotic leakage, pulmonary infections, and other severe complications but also prolongs hospitalization, escalates medical costs, and potentially impacts the long-term quality of life for patients.[7,8] Therefore, accurately predicting the duration of POGID in the early postoperative period holds significant importance in guiding personalized clinical interventions and enhancing patient prognosis.

At present, the prediction of POGID in esophageal cancer patients primarily relies on clinical experience and static single-factor indicators, such as age, comorbidities, and surgical duration.[9–11] However, these methods exhibit limitations in terms of low prediction accuracy and strong subjectivity. With the advancement of precision medicine and artificial intelligence technology, machine learning models have gained widespread application in the field of medical prediction, owing to their robust capability to handle high-dimensional complex data.[12,13] Nevertheless, most existing research has concentrated on predicting the occurrence of POGID rather than its specific duration.[9,14] Additionally, these studies typically utilize static data collected at a single time point, overlooking the dynamic changes in physiological and pathological indicators before and during the perioperative period. This oversight fails to fully reflect the body's response to surgery and the continuous process of POGID development.[11,15]

Compared to traditional static machine learning methods, time dynamic models such as LSTM can effectively capture the continuous time-dependent changes of physiological indicators during the perioperative period. This approach has rarely been reported in the field of predicting the duration of POGID in esophageal cancer patients. Blood routine examination is a simple and feasible clinical test that can reflect the body's inflammatory status, nutritional status, and hematopoietic function.[16] The dynamic changes of routine blood indicators (such as white blood cell count, C-reactive protein, and hemoglobin) before and during the perioperative period are closely related to the degree of surgical trauma, inflammatory response, and postoperative recovery. Ultrasound examination, as a non-invasive real-time imaging method, can dynamically assess the structure and function of the gastrointestinal tract, such as gastric sinus volume, intestinal peristalsis, and gallbladder contraction, directly reflecting the functional status of the gastrointestinal tract.[17] Therefore, integrating the temporal dynamic information of these two types of indicators can provide more comprehensive and accurate predictive information for the duration of POGID.

On this basis, this study aims to construct a temporal dynamic machine learning model by collecting preoperative and perioperative routine blood time data, ultrasound dynamic features, and clinical baseline data from cancer patients. By comparing the performance of dynamic and static models at different time points, we aim to select the best predictive model, identify key predictive features, and provide new methods and evidence for clinical prediction and intervention of POGID duration in cancer patients.

# Methods

## Study Population

This retrospective study was approved by the Ethics Committee of the First Affiliated Hospital of Xi'an Medical University. Due to the retrospective nature of the study, the requirement for informed consent was waived. The study population includes cancer patients with esophageal cancer who underwent radical resection (esophagectomy plus lymph node dissection) at the First Affiliated Hospital of Xi'an Medical University from January 2017 to December 2024. The inclusion criteria are as follows: (1) primary esophageal squamous cell carcinoma or adenocarcinoma confirmed by pathology; (2) Age between 18 and 80 years old; (3) First time undergoing radical surgical resection; (4) Complete clinical data, including preoperative and perioperative blood routine examination results, ultrasound examination data, and follow-up records of POGID duration; (5) Postoperative POGID (diagnosed according to the International Pancreatic Surgery Research Group's diagnostic criteria for gastrointestinal dysfunction: postoperative abdominal distension lasting for more than 24 hours, delayed recovery from oral feeding (inability to take solid food within 5 days after surgery), or the need for gastrointestinal decompression for more than 48 hours after surgery).[18] The exclusion criteria are: (1) combined with other malignant tumors; (2) serious preoperative gastrointestinal diseases (such as inflammatory bowel disease, gastric outlet obstruction); (3) severe functional impairment of important organs such as the liver, kidneys, and heart; (4) key data missing (blood routine indicators or ultrasound features missing by more than 30%); (5) death within 7 days after surgery (unable to assess the duration of POGID). The inclusion, screening, and analysis procedures for patients are shown in Figure 1.

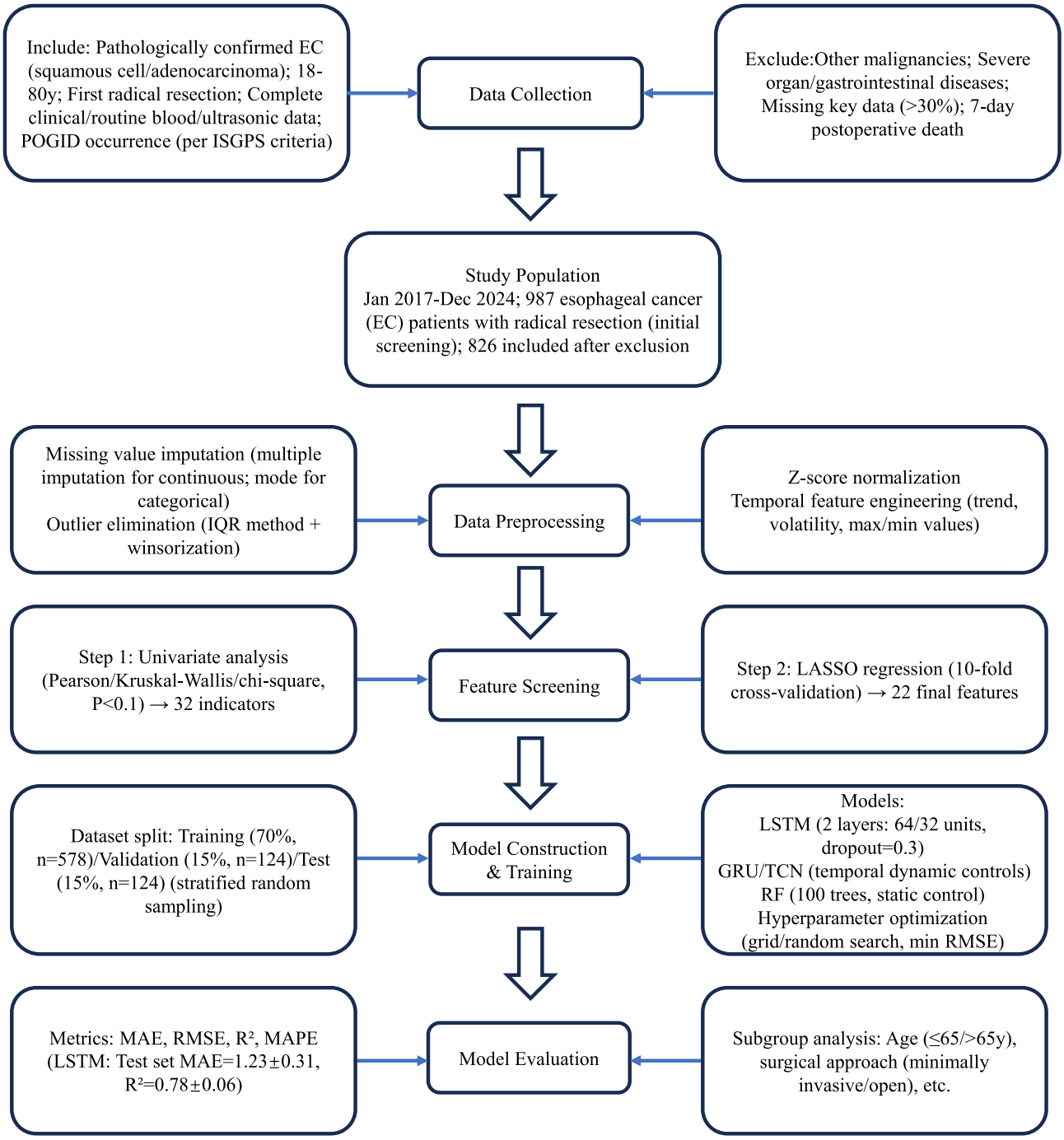


**Figure 1** Workflow of POGID predictive model in esophageal cancer post-radical resection.

## Data Collection and Indicator Screening

The data is sourced from the electronic medical record system, laboratory information system, and ultrasound examination information system of participating hospitals, encompassing three categories of indicators: clinical baseline data, preoperative and perioperative blood routine data, and ultrasound dynamic characteristics.

Clinical baseline data comprises: (1) demographic information: age, gender, body mass index (BMI); (2) comorbidities: hypertension (diagnosed according to the Chinese Guidelines for the Prevention and Treatment of Hypertension), diabetes (diagnosed according to the diagnostic criteria of the American Diabetes Association), coronary heart disease


https://doi.org/10.2147/IJGM.S588931

(diagnosed according to the diagnostic criteria of the World Health Organization); (3) tumor-related indicators: tumor location (upper, middle, lower esophagus), tumor size (maximum diameter measured by preoperative imaging), TNM staging (according to the 8th edition TNM staging system of the Joint Committee on Cancer in the United States); (4) surgical information: surgical approach (open surgery, minimally invasive surgery), surgical time, intraoperative blood loss, whether lymph node dissection was performed, and whether neoadjuvant therapy (chemotherapy or radiochemotherapy) was received before surgery.

Preoperative and perioperative blood routine data are collected at five fixed time points: 7 days before surgery (T1), 3 days before surgery (T2), 1 day before surgery (T3), 1 day after surgery (T4), and 3 days after surgery (T5). The indicators include white blood cell count (WBC), neutrophil count (NEUT), lymphocyte count (LYMPH), monocyte count (MONO), and eosinophil count. All indicators are tested by fully automated biochemical analyzers and blood analyzers in clinical laboratories of various hospitals, and the testing methods comply with national clinical laboratory quality standards.

Use a color Doppler ultrasound diagnostic instrument (model: GE Logiq E9, Philips EPIQ 7C) equipped with a 3.5–5.0 MHz convex array probe to collect ultrasound dynamic features at three time points: 1 day before surgery (U1), 2 days after surgery (U2), and 5 days after surgery (U3). The testing is conducted by an ultrasound doctor with over 5 years of work experience, and the results are reviewed by the attending physician. The indicators include: (1) gastric related features: cross-sectional area of the gastric antrum (measured in supine position, 3cm below the pylorus), thickness of the gastric wall (measured in the anterior wall of the gastric antrum), and frequency of gastric peristalsis (wave number per minute); (2) intestinal related characteristics: intestinal peristalsis frequency (number of waves per minute of ileum peristalsis), intestinal wall thickness (measured at the ileum wall), and intestinal diameter (maximum diameter of ileum); (3) gallbladder related characteristics: gallbladder volume (calculated according to ellipsoid formula: volume=0.5 × length × width × height), gallbladder wall thickness, gallbladder contraction rate (measured before and after eating, contraction rate=(pre meal volume - post meal volume)/pre meal volume × 100%); (4) pancreatic related features: thickness of pancreatic head, thickness of pancreatic body, and diameter of pancreatic duct.

The indicator screening process is divided into two steps: first, use univariate analysis to screen for indicators related to the duration of POGID. For continuous variables, calculate Pearson correlation coefficient, and for categorical variables, use Kruskal–Wallis *H*-test or chi-square test. Indicators with $P<0.1$ are included in the preliminary feature set. Secondly, the Least Absolute Shrinkage and Selection Operator (LASSO) regression is used to further screen features to eliminate multicollinearity.[19] The optimal lambda value is determined through 10-fold cross-validation, and indicators with non-zero coefficients are included in the final feature set of the model construction.[20]

## Prediction Model Construction

We utilized stratified random sampling to partition the study population into a training set (70%), a validation set (15%), and a testing set (15%). Data preprocessing prior to model construction includes the following steps: (1) Missing value imputation: For missing values in continuous variables (with a missing rate of less than 30%), multiple imputation methods are employed. For categorical variables, this specific pattern is utilized for imputation.[21,22] Variables with a missing rate exceeding 30% are excluded. (2) Outlier elimination: The interquartile range method is applied to identify outliers (IQR=Q3-Q1, values less than Q1-1.5×IQR or greater than Q3+1.5×IQR are considered outliers). The winsorization method is then utilized to handle outliers (replacing them with Q1-1.5×IQR or Q3+1.5×IQR, respectively). (3) Normalization: The Z-score normalization technique is adopted to normalize continuous variables, thereby eliminating the impact of dimensionality. (4) Feature engineering: For time-series data, such as blood routine examination results and ultrasound features, we construct time series features including trends (slope of indicator change curve), volatility (standard deviation of indicators at various time points), and maximum/minimum values to enhance the representation of temporal information.

Four machine learning models were constructed: (1) LSTM: input layer, two LSTM layers (64 and 32 hidden units), dropout layer (0.3), and fully connected output layer. LSTM is a recurrent neural network that utilizes input, forget, and output gates to capture long-term dependencies in time series data and prevent gradient vanishing.[23] (2) GRU: a simplified version of LSTM that merges the forget and input gates into an update gate, thereby reducing the number of parameters. The model structure mirrors that of LSTM, featuring two GRU layers (with 64 and 32 hidden units

respectively).[24] (3) TCN: a convolutional neural network model equipped with extended convolution, capable of capturing multi-scale temporal information. This model comprises four TCN layers (with 32, 64, 64, and 32 filters respectively) and a fully connected output layer.[25] (4) RF: a traditional machine learning model based on decision trees, serving as a static control model. This model employs static features (average of time indicators) and clinical baseline data for prediction, incorporating 100 decision trees.[26] We trained the model using a training set and optimized the hyperparameters using a validation set. Optimization methods included grid search and random search, with the evaluation metric being the minimum Root Mean Squared Error (RMSE). The test set was used to evaluate the performance of the final model.

## Statistical Methods

We utilized Python 3.9 (with Scikit-learn, TensorFlow, and Keras libraries) and SPSS 26.0 software for statistical analysis. Depending on the distribution type, measurement data is presented as mean ± standard deviation (SD) or median (interquartile range), while count data is represented as frequency (percentage). We compared the performance of various model groups using paired t-tests or Wilcoxon signed rank tests. Pearson or Spearman correlation coefficients were employed to analyze the correlation between indicators and the duration of POGID. Receiver operating characteristic (ROC) curves were used to assess the model's classification performance, specifically when POGID duration was categorized into short (≤5 days) and long (>5 days) durations, and the area under the curve (AUC) was calculated. SHAP values were utilized to analyze feature importance, and SHAP summary plots were employed to visualize the contribution of each feature to model prediction. A two-tailed P-value < 0.05 was considered statistically significant.

# Results

## Baseline Characteristics of the Study Population

We initially screened 987 patients with esophageal cancer who underwent radical resection. After excluding 161 patients who did not meet the inclusion criteria or had missing key data, we ultimately included 826 patients in the study. The research population was divided into a training set (n=578), a validation set (n=124), and a testing set (n=124). The median duration of POGID in the overall population was 6.2 (4.1–8.5) days. As shown in Table 1 and Supplementary Figure 1, there were no significant differences in age, gender, BMI, comorbidities (hypertension, diabetes, coronary heart disease), tumor

**Table 1** Baseline Characteristics of Esophageal Cancer Patients Across All Datasets

| Characteristic | Training Set (n=578) | Validation Set (n=124) | Test Set (n=124) | P value | Test Method |
|---|---|---|---|---|---|
| **Continuous Variables** | | | | | |
| Age (years) | 60.9±8.2 | 61.3±7.8 | 60.7±8.0 | 0.815 | ANOVA |
| BMI (kg/m$^2$) | 23.2±3.1 | 22.9±3.3 | 23.4±3.0 | 0.672 | ANOVA |
| Tumor size (cm) | 4.3(3.2–5.4) | 4.2(3.1–5.3) | 4.4(3.3–5.5) | 0.738 | Kruskal–Wallis |
| Operation time (min) | 194.6±32.8 | 193.2±31.7 | 195.8±33.0 | 0.601 | ANOVA |
| Intraoperative blood loss (mL) | 151.8±45.5 | 149.3±43.8 | 153.2±46.1 | 0.647 | ANOVA |
| POGID duration (days) | 6.1(4.0–8.4) | 6.3(4.2–8.6) | 6.2(4.1–8.5) | 0.912 | Kruskal–Wallis |
| Preoperative albumin (g/L) | 38.5±3.6 | 38.2±3.4 | 38.7±3.5 | 0.703 | ANOVA |
| Preoperative hemoglobin (g/L) | 135.2±15.8 | 134.6±15.3 | 136.1±15.5 | 0.789 | ANOVA |
| Postop Day1 WBC (×10$^9$/L) | 10.2(8.5–12.1) | 10.4(8.7–12.3) | 10.1(8.4–11.9) | 0.856 | Kruskal–Wallis |
| **Categorical Variables, n (%)** | | | | | |
| Gender | | | | 0.802 | Chi-square Test |
| Male | 440 (76.1) | 96 (77.4) | 93 (75.0) | | |
| Female | 138 (23.9) | 28 (22.6) | 31 (25.0) | | |
| Hypertension | | | | 0.887 | Chi-square Test |
| Yes | 200 (34.6) | 42 (33.9) | 43 (34.7) | | |
| No | 378 (65.4) | 82 (66.1) | 81 (65.3) | | |

(*Continued*)


https://doi.org/10.2147/IJGM.S588931

**Table 1** (Continued).

| Characteristic | Training Set (n=578) | Validation Set (n=124) | Test Set (n=124) | P value | Test Method |
|---|---|---|---|---|---|
| Diabetes | | | | 0.935 | Chi-square Test |
| Yes | 88 (15.2) | 18 (14.5) | 19 (15.3) | | |
| No | 490 (84.8) | 106 (85.5) | 105 (84.7) | | |
| Coronary heart disease | | | | 0.768 | Chi-square Test |
| Yes | 62 (10.7) | 13 (10.5) | 14 (11.3) | | |
| No | 516 (89.3) | 111 (89.5) | 110 (88.7) | | |
| COPD | | | | 0.824 | Chi-square Test |
| Yes | 45 (7.8) | 10 (8.1) | 9 (7.3) | | |
| No | 533 (92.2) | 114 (91.9) | 115 (92.7) | | |
| Tumor location | | | | 0.836 | Chi-square Test |
| Upper thoracic | 91 (15.7) | 21 (16.9) | 18 (14.5) | | |
| Middle thoracic | 329 (56.9) | 69 (55.6) | 73 (58.9) | | |
| Lower thoracic | 158 (27.3) | 34 (27.4) | 33 (26.6) | | |
| Tumor differentiation | | | | 0.901 | Chi-square Test |
| Well | 128 (22.1) | 27 (21.8) | 26 (21.0) | | |
| Moderate | 312 (53.9) | 68 (54.8) | 67 (54.0) | | |
| Poor | 138 (23.9) | 29 (23.4) | 31 (25.0) | | |
| TNM stage | | | | 0.942 | Chi-square Test |
| I | 85 (14.7) | 18 (14.5) | 19 (15.3) | | |
| II | 271 (46.9) | 58 (46.8) | 59 (47.6) | | |
| III | 202 (34.9) | 43 (34.7) | 42 (33.9) | | |
| IV | 20 (3.5) | 5 (4.0) | 4 (3.2) | | |
| Surgical approach | | | | 0.875 | Chi-square Test |
| Thoracoscopic | 386 (66.8) | 81 (65.3) | 84 (67.7) | | |
| Open | 192 (33.2) | 43 (34.7) | 40 (32.3) | | |
| Neoadjuvant therapy | | | | 0.813 | Chi-square Test |
| Yes | 125 (21.6) | 26 (21.0) | 28 (22.6) | | |
| No | 553 (78.4) | 98 (79.0) | 96 (77.4) | | |
| Postoperative complication | | | | 0.926 | Chi-square Test |
| Yes | 116 (20.1) | 25 (20.2) | 24 (19.4) | | |
| No | 462 (79.9) | 99 (79.8) | 100 (80.6) | | |

**Abbreviations**: BMI, Body mass index; POGID, Postoperative gastrointestinal dysfunction; WBC, White blood cell count; COPD, Chronic obstructive pulmonary disease.

location, tumor size, TNM stage, surgical approach, surgical time, intraoperative blood loss, and neoadjuvant treatment among the three groups of patients (all P>0.05). This indicates that the grouping was reasonable and the data were comparable. The baseline characteristics of the study population were balanced across different datasets, laying a reliable foundation for subsequent model construction and performance evaluation.

## Feature Screening Results

Through univariate analysis, we included a total of 56 initial indicators and selected 32 indicators with P<0.1 (Supplementary Table 1). After LASSO regression analysis, the model construction ultimately incorporated 22 indicators, including 5 clinical baseline indicators (age, TNM stage, surgical approach, surgical time, and whether new adjuvant therapy was received), 11 routine blood time indicators (T3 WBC, T4 WBC, T5 WBC, T3 Hb, T4 PLT, T5 PLT, T3 CRP, T4 CRP, T4 ALB, and T3 NLR), and 6 ultrasound dynamic features (U1 antral cross-sectional area, U2 antral cross-sectional area, U2 intestinal peristalsis frequency, U3 intestinal peristalsis frequency, U1 gallbladder volume, and U2 gallbladder wall thickness). The distribution of LASSO regression coefficients and cross-validation curves are shown in Figure 2A and B, which confirms that the selected features exhibit good discriminative ability. The feature selection process effectively eliminated irrelevant and redundant indicators, reducing the complexity of the model while retaining key information related to the duration of POGID.

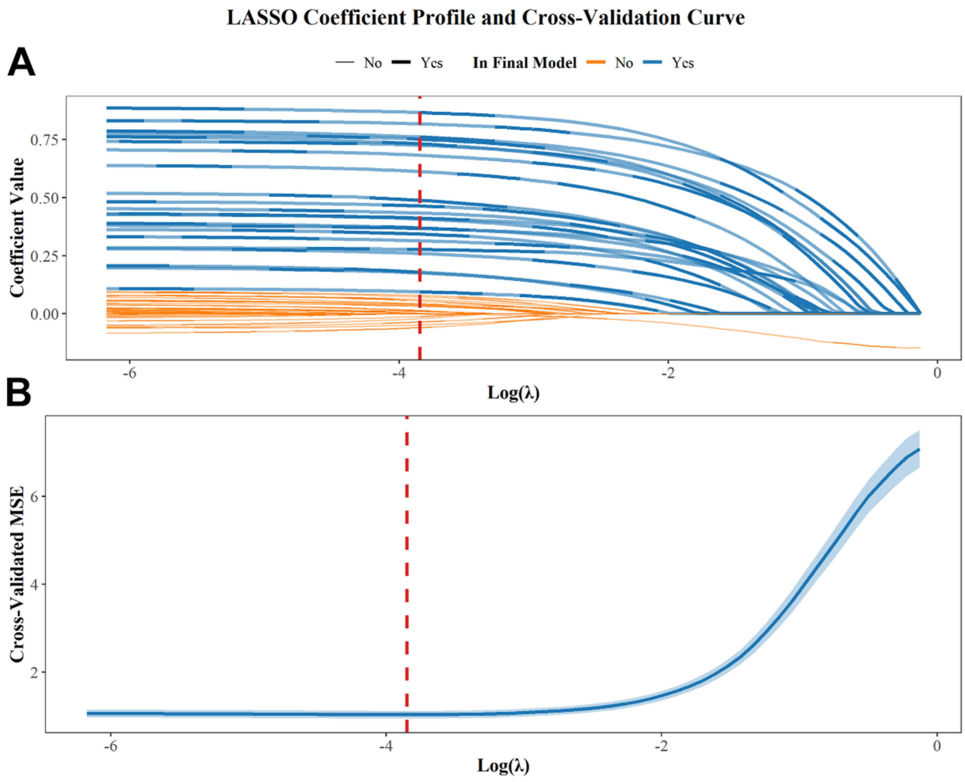


**Figure 2** LASSO coefficient profile (**A**) and cross-validation curve (**B**) for POGID predictive model construction.

## Performance Comparison of Different Prediction Models

The performance of the four models in the training set, validation set, and test set is illustrated in Figure 3A and B. In the training set, all models exhibited good fitting performance, yet the LSTM model demonstrated the lowest MAE (0.98 ± 0.25 days) and RMSE (1.21 ± 0.33 days), along with the highest $R^2$ (0.85 ± 0.05). In the validation set, despite a slight decline in performance across all models, the LSTM model maintained its optimal performance (MAE: 1.12 ± 0.28 days, RMSE: 1.42 ± 0.38 days, $R^2$: 0.81 ± 0.06). In the test set, the LSTM model continued to exhibit superior predictive performance, achieving a MAE of 1.23 ± 0.31 days, RMSE of 1.56 ± 0.42 days, $R^2$ of 0.78 ± 0.06, and MAPE of 12.3% ± 3.1%. When compared to the GRU model (MAE: 1.45 ± 0.35 days, RMSE: 1.82 ± 0.45 days, $R^2$: 0.71 ± 0.07, MAPE: 14.8% ± 3.5%), TCN model (MAE/1.58 ± 0.38 days, RMSE: 1.95 ± 0.48 days, $R^3$: 0.68 ± 0.08, MAPE: 16.2% ± 3.8%), and RF model (MAE-2.15 ± 0.45 days, RMSE: 2.68 ± 0.53 days, $R^2$: 0.52 ± 0.08, MAPE: 21.5% ± 4.2%), the LSTM model significantly outperformed them (all $P < 0.001$). The superior performance of the LSTM model over GRU and TCN can primarily be attributed to its enhanced ability to capture long-term temporal dependencies in perioperative sequence data. Conversely, GRU is a simplified variant with diminished feature extraction capabilities, while TCN faces limitations in modeling long sequence data. The distribution of prediction errors for the LSTM model in the test set is depicted in Figure 3C, with the majority of errors falling within ±2 days, indicating the model's high practical application value. The time-dynamic LSTM model surpasses other models, affirming the significance of integrating temporal data for predicting the duration of POGID.

## Feature Importance Analysis Results

The SHAP value analysis results of the LSTM model are presented in Figure 4A and B. The top five significant features influencing the duration of POGID are T4 white blood cells (SHAP value: 0.86), U1 gastric antral cross-sectional area (SHAP value: 0.79), T5 PLT (SHAP value: 0.72), TNM staging (SHAP score: 0.68), and U2 intestinal peristalsis frequency (SHAP ratio: 0.65). Among these, T4 white blood cells (positively correlated, r=0.62) and TNM staging (positively correlated, r=0.58) are positively associated with the duration of POGID, suggesting that higher levels of T4 white blood cells and later TNM stages are related to longer POGID durations. Conversely, the U1 gastric antral cross-


https://doi.org/10.2147/IJGM.S588931

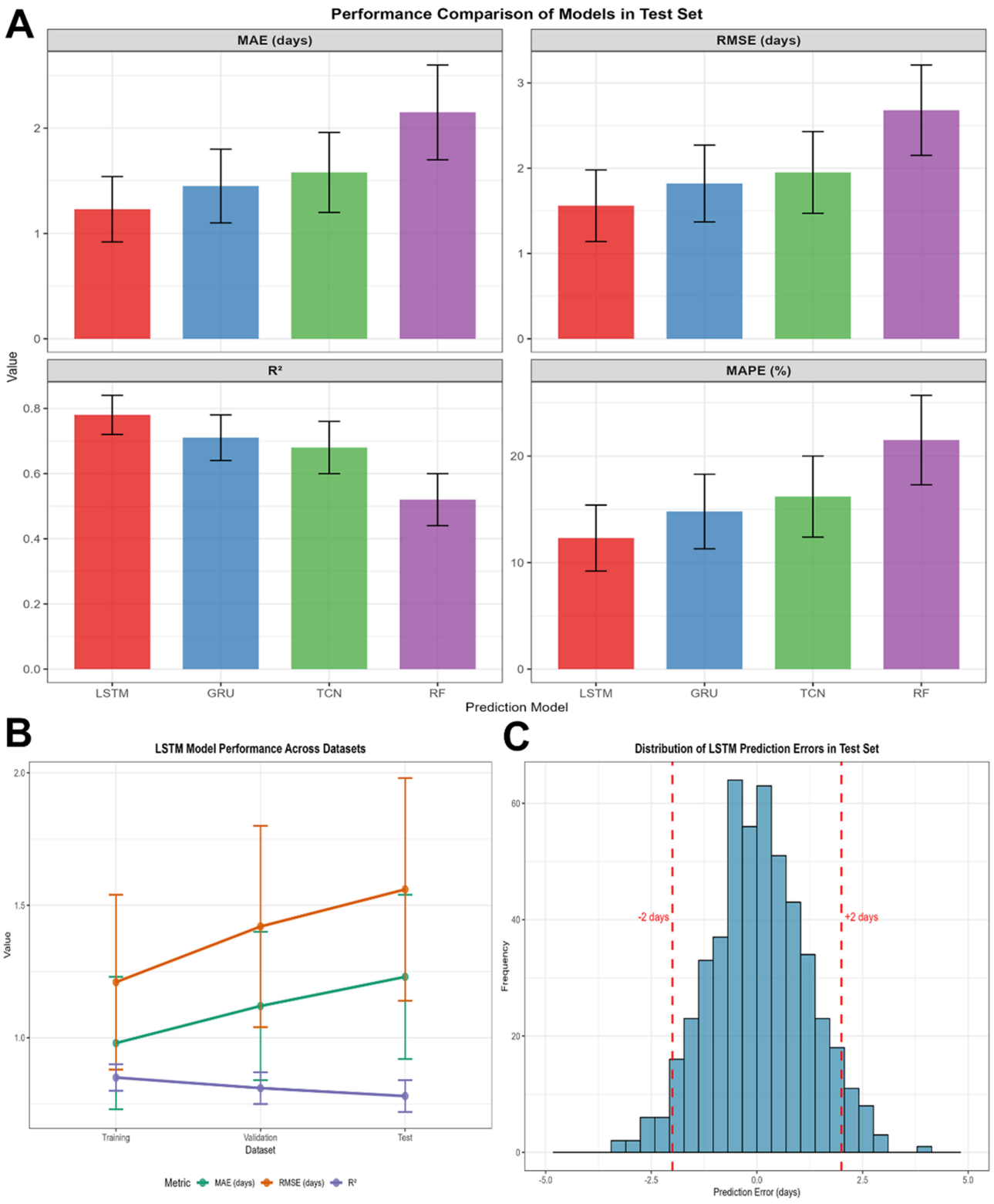


**Figure 3** Performance evaluation of prediction models for POGID duration (**A** and **B**) and prediction error distribution of LSTM model (**C**).

sectional area (negatively correlated, r=−0.55), T5 PLT (negatively correlated, r=−0.51), and U2 intestinal peristalsis frequency (negatively correlated, r=−0.49) have a negative impact on POGID duration. That is, a smaller U1 gastric antral cross-sectional area, higher T5 PLT, and increased U2 intestinal peristalsis frequency are associated with shorter


https://doi.org/10.2147/IJGM.S588931

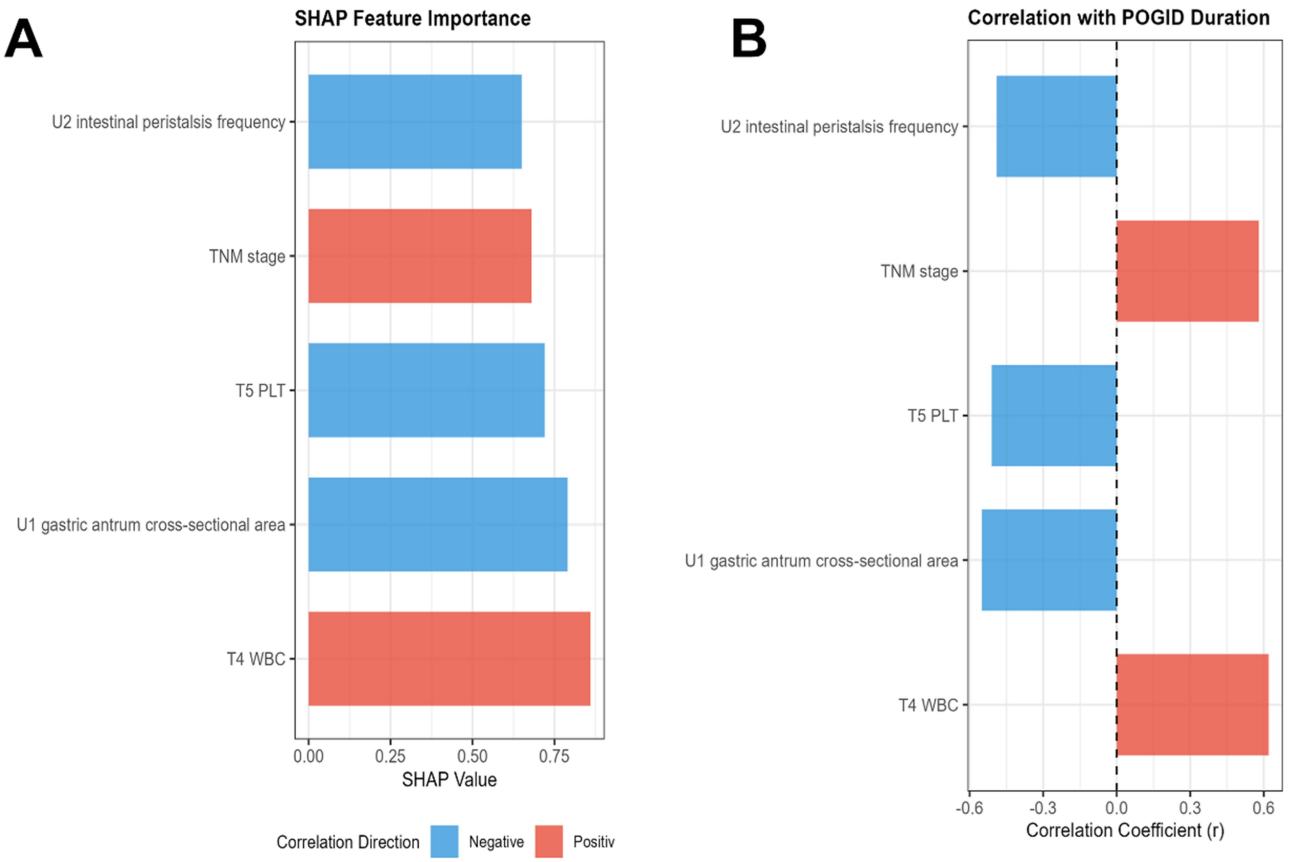


**Figure 4** SHAP summary plot (**A**) shows the contribution degree of each feature to model prediction; (**B**) shows correlation analysis between key features and POGID duration.

POGID durations. The key predictive features identified through SHAP analysis offer clear objectives for clinical interventions, facilitating the development of targeted prevention and treatment strategies.

## Subgroup Analysis Results

Subgroup analysis was conducted based on age (≤65 years and >65 years), surgical approach (minimally invasive surgery and open surgery), and comorbidity status (with and without comorbidities). The results are presented in Figure 5A–D. In the subgroup aged ≤65 years, the $R^2$ of the LSTM model is 0.81 ± 0.07; in the subgroup aged >65 years, the $R^2$ is 0.76 ± 0.08, with no significant difference observed between the two subgroups (P=0.12). In the minimally invasive surgery subgroup, the $R^2$ is 0.83 ± 0.06; in the open surgery subgroup, the $R^2$ is 0.72 ± 0.09, indicating a statistically significant difference (P=0.03), yet the $R^2$ remains above 0.7. In the subgroup without comorbidities, the $R^2$ is 0.80 ± 0.07; in the subgroup with comorbidities, the $R^2$ is 0.75 ± 0.08, with no statistically significant difference noted (P=0.09). These results suggest that the LSTM model exhibits stable predictive performance across different subgroups. The consistent performance of the LSTM model in various subgroups underscores its applicability in diverse clinical settings.

# Discussion

Our study constructed a temporal dynamic LSTM model to predict the duration of POGID in cancer patients by integrating preoperative and perioperative routine blood time data, ultrasound dynamic features, and clinical baseline data. The results indicate that the LSTM model exhibits excellent predictive performance, with an $R^2$ of 0.78 in the test set, significantly surpassing the static RF model and other time dynamic models. Postoperative inflammatory response inhibits gastrointestinal motility, which constitutes the core pathophysiological mechanism underlying the prolongation of POGID. On the first day after surgery, white blood cell count reflects the intensity of acute inflammatory response,


https://doi.org/10.2147/IJGM.S588931

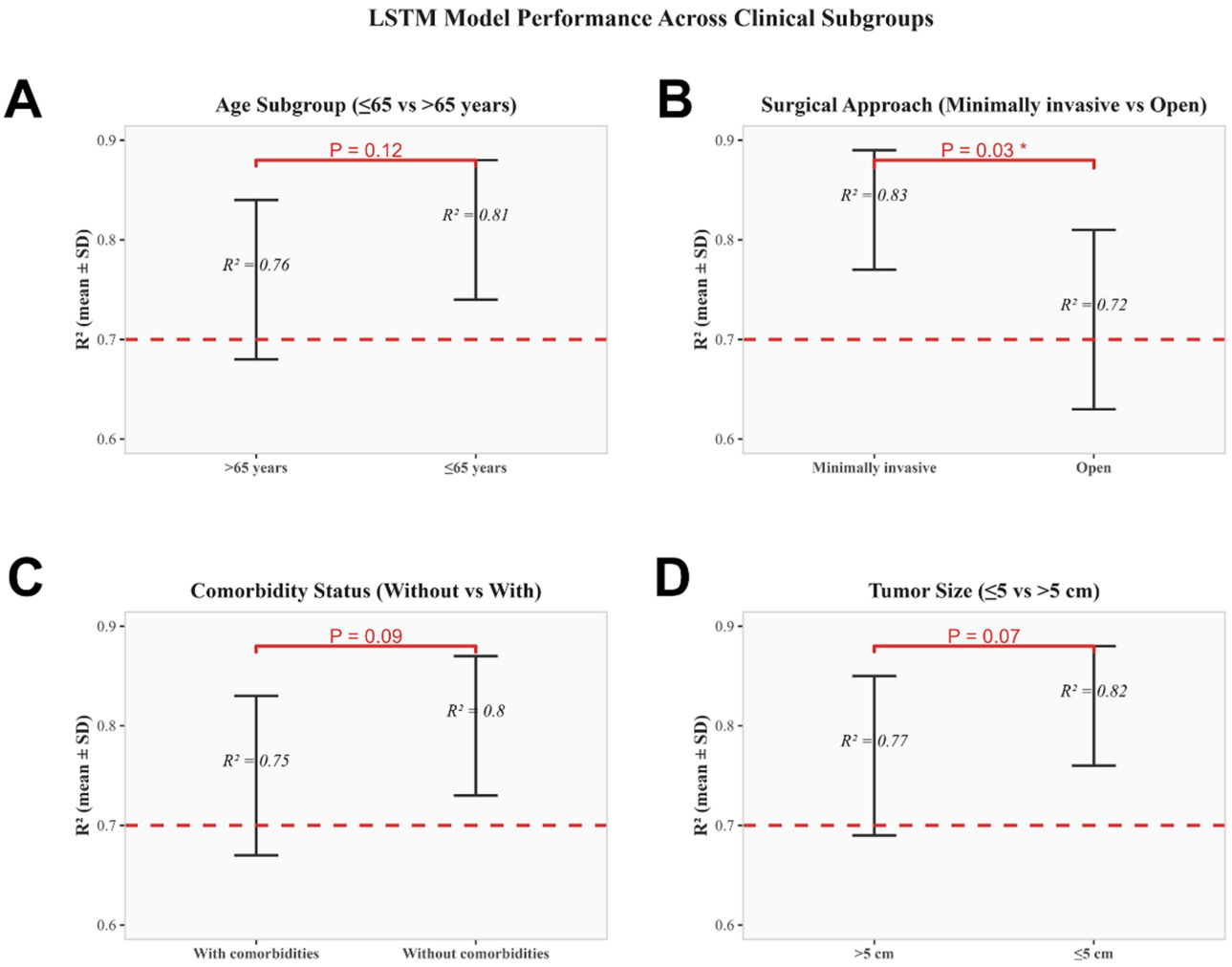


**Figure 5** LSTM model performance in age (**A**), surgical approach (**B**), comorbidity (**C**), and Tumor size(**D**) subgroups. Asterisk (*) indicates a statistically significant difference (P = 0.03, P < 0.05) between subgroups.

while the preoperative gastric antral cross-sectional area indicates baseline gastric functional reserve. Postoperative platelet elevation facilitates mucosal repair and inflammation resolution. Late TNM staging signifies more severe systemic stress and greater surgical trauma, and postoperative intestinal peristalsis frequency directly mirrors the recovery of gastrointestinal function. This model can be integrated into clinical decision support systems to facilitate early risk stratification and precise intervention.

The advantage of the time dynamic model lies in its ability to capture the dynamic changes of physiological and pathological indicators during the preoperative and perioperative periods, aligning more closely with the clinical reality of the continuous development of POGID. Previous studies on the prediction of POGD in gynecological cancer primarily concentrated on its occurrence. For instance, He et al constructed logistic regression models using static clinical indicators (preoperative serum bilirubin and sodium levels) and intraoperative factors (surgical duration) to predict the occurrence of POGD. However, these predictions only highlighted high-risk factors without providing specific details on the severity or duration of POGD, potentially limiting the refinement of personalized perioperative intervention plans.[10] Some studies have explored perioperative complications in elderly patients with gastrointestinal diseases, but they either focused on static baseline indicators or relied on discrete single time point assessments. For example, Yoshiyama et al utilized preoperative factors such as age, sarcopenia, and MMSE score to construct a logistic regression model for predicting postoperative swallowing dysfunction in elderly gastrointestinal cancer patients, elucidating independent risk factors, but failed to explore the dynamic recovery trajectory of this dysfunction. Potestio et al analyzed the correlation

between EEG (PSI) indicators processed during propofol sedation and postoperative cognitive impairment (POCD) in elderly patients undergoing gastrointestinal endoscopy through short-term blessing tests (SBT) at 7 and 90 days postoperatively. However, they found no significant correlation between sedation depth and cognitive changes, and there was also a lack of dynamic tracking of perioperative cognitive fluctuations.[11,27] These limitations hinder the development of detailed, stage-specific clinical intervention plans. Compared to these studies, our research not only focuses on the duration of POGID but also integrates time data from multiple time points, significantly improving prediction accuracy. This improvement may be attributed to the fact that time data can reflect the dynamic process of the body's response to surgical trauma, such as the increase in white blood cells after surgery indicating an inflammatory response, and the recovery of PLT reflecting improved coagulation function and tissue repair ability, all of which are closely related to the recovery of gastrointestinal function.

The analysis of feature importance reveals that white blood cells on the first postoperative day are the most significant predictive feature. This aligns with clinical insights that inflammatory response is a crucial factor influencing postoperative gastrointestinal function. Surgical trauma triggers the activation of inflammatory cells and the release of inflammatory factors, subsequently inhibiting gastrointestinal motility and delaying the recovery of gastrointestinal function.[28,29] A study conducted by Kampman et al revealed that, compared to open surgery, laparoscopic cancer surgery is associated with lower levels of IL-6 and CRP on the first postoperative day, as well as lower CRP levels on the third postoperative day.[30] However, there was no difference in the incidence of postoperative infection complications, providing evidence for the superior inflammatory characteristics of laparoscopic surgery. Preoperative ultrasound measurement of the pyloric antral cross-sectional area (CSA) serves as another crucial indicator for assessing high-risk gastric status. The gastric antral CSA, measured in a supine position, reflects the volume of gastric contents. A larger gastric antral CSA increases the risk of solid/viscous liquid gastric contents or a gastric volume exceeding 1.5 mL/kg. Okada et al's study supports this notion, reporting a sensitivity of 85% and a negative predictive value of 53% for a critical gastric antral CSA value of 3.01 $cm^2$, used to identify "high-risk stomach". Furthermore, the correlation between gastric antral CSA and CT-calculated gastric volume is enhanced, particularly in stomachs containing solid contents/viscous liquids.[31]

On the third day after surgery, there was a negative correlation between PLT and POGID duration. Platelets not only play a role in coagulation but also participate in tissue repair and inflammation regulation. After surgery, an increase in PLT can promote the repair of intestinal mucosa and the resolution of inflammatory reactions, thereby accelerating the recovery of gastrointestinal function. A study by Weng et al confirmed that early postoperative enteral nutrition (EPEN) is associated with better postoperative recovery in patients undergoing gastrointestinal tumor surgery, which is consistent with our findings.[32] Tumor TNM staging is also an important predictive feature, possibly because advanced tumors are often accompanied by more severe systemic effects and stronger surgical trauma, resulting in a longer recovery time for gastrointestinal function. The frequency of intestinal peristalsis on the second day after surgery directly reflects the recovery of intestinal movement and therefore has significant predictive value for the duration of POGID.

The LSTM model constructed in this study exhibits stable performance across various subgroups, including different age groups, surgical methods, and comorbidity states. Particularly within the minimally invasive surgery subgroup, the model demonstrates high predictive accuracy. This may be attributed to the minimal trauma associated with minimally invasive surgery, more stable intraoperative hemodynamics, and more regular changes in time indicators, all of which facilitate model learning. Although the prediction accuracy in the open surgery subgroup is slightly lower, the $R^2$ remains above 0.7, indicating that the model still holds significant application value. For elderly patients and those with comorbidities, who constitute high-risk groups for POGID, the model maintains stable predictive performance, which is crucial for clinical targeted interventions.

The research we conducted also inevitably has some limitations. Firstly, this is a retrospective single-center study with potential selection bias. Secondly, external prospective validation is needed to improve generalization ability. Thirdly, the data comes from three tertiary hospitals, and the results may not be applicable to grassroots hospitals with different medical conditions. Fourthly, the model has not been validated in a prospective cohort, and its clinical efficacy needs further confirmation. In the future, we will conduct prospective studies to validate the model, expand the sample size and the scope of participating hospitals, and further improve the model's generalization ability.

## Conclusion

In summary, we integrated preoperative and perioperative routine blood time data, ultrasound dynamic features, and clinical baseline data based on a time dynamic LSTM model, which can accurately predict the duration of POGID in cancer patients. The key predictive features provide potential targets for clinical interventions. This model is expected to be applied in clinical decision support systems to optimize perioperative management and improve postoperative recovery outcomes. Further prospective validation is needed.

## Disclosure

The authors report no conflicts of interest in this work.

## References


1. Jiang W, Zhang B, Xu J, Xue L, Wang L. Current status and perspectives of esophageal cancer: a comprehensive review. *Cancer Commun*. 2025;45(3):281–331. doi:10.1002/cac2.12645
2. Uhlenhopp DJ, Then EO, Sunkara T, Gaduputi V. Epidemiology of esophageal cancer: update in global trends, etiology and risk factors. *Clin J Gastroenterol*. 2020;13(6):1010–1021. doi:10.1007/s12328-020-01237-x
3. Huang FL, Yu SJ. Esophageal cancer: risk factors, genetic association, and treatment. *Asian J Surg*. 2018;41(3):210–215. doi:10.1016/j.asjsur.2016.10.005
4. Nakamura M, Kido Y, Hosoya Y, Yano M, Nagai H, Monden M. Postoperative gastrointestinal dysfunction after 2-field versus 3-field lymph node dissection in patients with esophageal cancer. *Surg Today*. 2007;37(5):379–382. doi:10.1007/s00595-006-3413-4
5. Loan BTH, Nakahara S, Tho BA, et al. Nutritional status and postoperative outcomes in patients with gastrointestinal cancer in Vietnam: a retrospective cohort study. *Nutrition*. 2018;48:117–121. doi:10.1016/j.nut.2017.11.027
6. Wan J, Che Y, Kang N, Zhang R. Surgical method, postoperative complications, and gastrointestinal motility of thoraco-laparoscopy 3-field esophagectomy in treatment of esophageal cancer. Med Sci Monit. 2016;22:2056–2065. doi:10.12659/MSM.895882
7. Saito S, Nakamura M, Hosoya Y, Kitayama J, Lefor AK, Sata N. Postoperative quality of life and dysfunction in patients after combined total gastrectomy and esophagectomy. *Ann Med Surg*. 2017;22:34–38. doi:10.1016/j.amsu.2017.08.016
8. Haverkort EB, Binnekade JM, Busch OR, van Berge Henegouwen MI, de Haan RJ, Gouma DJ. Presence and persistence of nutrition-related symptoms during the first year following esophagectomy with gastric tube reconstruction in clinically disease-free patients. *World J Surg*. 2010;34(12):2844–2852. doi:10.1007/s00268-010-0786-8
9. Huang H, Chou J, Tang Y, Ouyang W, Wu X, Le Y. Nomogram to predict postoperative cognitive dysfunction in elderly patients undergoing gastrointestinal tumor resection. *Front Aging Neurosci*. 2022;14:1037852. doi:10.3389/fnagi.2022.1037852
10. He L, Hu J, Han Y, Xiong W. Predictive modeling of postoperative gastrointestinal dysfunction: the role of serum bilirubin, sodium levels, and surgical duration in gynecological cancer care. *BMC Women's Health*. 2023;23(1):598. doi:10.1186/s12905-023-02779-1
11. Yoshiyama S, Kawabata R, Yamaguchi H, et al. Risk factors for postoperative swallowing dysfunction in elderly patients undergoing gastrointestinal cancer surgery: a retrospective cohort study. *Surg Today*. 2025. doi:10.1007/s00595-025-03185-w
12. Xiang Y, Zhao L, Liu Z, et al. Implementation of artificial intelligence in medicine: status analysis and development suggestions. *Artif Intell Med*. 2020;102:101780. doi:10.1016/j.artmed.2019.101780
13. Reel PS, Reel S, Pearson E, Trucco E, Jefferson E. Using machine learning approaches for multi-omics data analysis: a review. *Biotechnol Adv*. 2021;49:107739. doi:10.1016/j.biotechadv.2021.107739
14. Wang XH, Wang Y, Liu X, Ouyang XL, Xie F, Liao L. Construction of a risk prediction model for postoperative gastrointestinal dysfunction and prevention in patients with gastrointestinal tumors. *World J Gastrointest Surg*. 2026;18(2):114483. doi:10.4240/wjgs.v18.i2.114483
15. Qin Q, Huang B, Wu A, et al. Development and validation of a post-radiotherapy prediction model for bowel dysfunction after rectal cancer resection. *Gastroenterology*. 2023;165(6):1430–42.e14. doi:10.1053/j.gastro.2023.08.022
16. Han Y, Zheng S, Chen Y. Prognostic value of lymphocyte to monocyte ratio in patients with esophageal cancer: a systematic review and meta-analysis. *Front Oncol*. 2024;14:1401076. doi:10.3389/fonc.2024.1401076
17. Xia L, Sun S, Dai W. Deep learning-based ultrasound combined with gastroscope for the diagnosis of upper gastrointestinal submucous lesions. *Comput Math Methods Med*. 2022;2022:1607099. doi:10.1155/2022/1607099
18. Mazzotta E, Villalobos-Hernandez EC, Fiorda-Diaz J, Harzman A, Christofi FL. Postoperative ileus and postoperative gastrointestinal tract dysfunction: pathogenic mechanisms and novel treatment strategies beyond colorectal enhanced recovery after surgery protocols. *Front Pharmacol*. 2020;11:583422. doi:10.3389/fphar.2020.583422
19. Buch G, Schulz A, Schmidtmann I, Strauch K, Wild PS. A systematic review and evaluation of statistical methods for group variable selection. *Stat Med*. 2023;42(3):331–352. doi:10.1002/sim.9620
20. Mohr F, van Rijn JN. Fast and Informal Model Selection Using Learning Curve Cross-Validation. *IEEE Trans Pattern Anal Mach Intell*. 2023;45(8):9669–9680. doi:10.1109/TPAMI.2023.3251957
21. Austin PC, White IR, Lee DS, van Buuren S. Missing data in clinical research: a tutorial on multiple imputation. *Canad J Cardiol*. 2021;37(9):1322–1331. doi:10.1016/j.cjca.2020.11.010
22. Xu T, Chen K, Li G. The more data, the better? Demystifying deletion-based methods in linear regression with missing data. *Stat Interface*. 2022;15(4):515–526. doi:10.4310/21-SII717
23. Wang J, Liu S. Visual information computing and processing model based on artificial neural network. *Comput Intell Neurosci*. 2022;2022:4713311. doi:10.1155/2022/4713311
24. Li W, Situ Y, Ding L, Chen Y, Yang Q. MOF-GRU: a MOFid-aided deep learning model for predicting the gas separation performance of metal-organic frameworks. *ACS Appl Mater Interfaces*. 2023;15(51):59887–59894. doi:10.1021/acsami.3c11790

25. Li W, Wei Y, An D, Jiao Y, Wei Q. LSTM-TCN: dissolved oxygen prediction in aquaculture, based on combined model of long short-term memory network and temporal convolutional network. *Environ Sci Pollut Res Int*. 2022;29(26):39545–39556. doi:10.1007/s11356-022-18914-8
26. Trzepieciński T, Najm SM, Ibrahim OM, Kowalik M. Analysis of the frictional performance of AW-5251 aluminium alloy sheets using the random forest machine learning algorithm and multilayer perceptron. *Materials*. 2023;16(15). doi:10.3390/ma16155207
27. Potestio CP, Dibato J, Bolkus K, et al. Post-operative cognitive dysfunction in elderly patients receiving propofol sedation for gastrointestinal endoscopies. *Cureus*. 2023;15(10):e46588. doi:10.7759/cureus.46588
28. Bolte LA, Vich Vila A, Imhann F, et al. Long-term dietary patterns are associated with pro-inflammatory and anti-inflammatory features of the gut microbiome. *Gut*. 2021;70(7):1287–1298. doi:10.1136/gutjnl-2020-322670
29. Koirala U, Thapa PB, Joshi MR, Singh DR, Sharma SK. Systemic inflammatory response syndrome following gastrointestinal surgery. *JNMA*. 2017;56(206):221–225.
30. Kampman SL, Smalbroek BP, Dijksman LM, Smits AB. Postoperative inflammatory response in colorectal cancer surgery: a meta-analysis. *Int J Colorectal Dis*. 2023;38(1):233. doi:10.1007/s00384-023-04525-3
31. Okada Y, Toyama H, Kamata K, Yamauchi M. A clinical study comparing ultrasound-measured pyloric antrum cross-sectional area to computed tomography-measured gastric content volume to detect high-risk stomach in supine patients undergoing emergency abdominal surgery. *J Clin Monit Comput*. 2020;34(5):875–881. doi:10.1007/s10877-019-00438-1
32. Weng PH, Du Y, Yuan F. Early postoperative enteral nutrition and its impact on postoperative recovery in gastrointestinal tumor surgery. *Front Oncol*. 2025;15:1649306. doi:10.3389/fonc.2025.1649306