**Dovepress**

Open Access Full Text Article

ORIGINAL RESEARCH

# Identification of High-Risk Patients for Postoperative Myocardial Injury After CME Using Machine Learning: A 10-Year Multicenter Retrospective Study

Yuan Liu [iD]*, Chen Song*, Zhiqiang Tian, Wei Shen

Department of General Surgery, The Affiliated Wuxi People's Hospital of Nanjing Medical University, Wuxi, People's Republic of China

*These authors contributed equally to this work

Correspondence: Wei Shen, Department of General Surgery, The Affiliated Wuxi People's Hospital of Nanjing Medical University, Wuxi, Jiangsu Province, 214000, People's Republic of China, Tel +86 13385110723, Email shenweijs@outlook.com

**Purpose:** The occurrence of myocardial injury, a grave complication post complete mesocolic excision (CME), profoundly impacts the immediate and long-term prognosis of patients. The aim of this inquiry was to conceive a machine learning model that can recognize preoperative, intraoperative and postoperative high-risk factors and predict the onset of myocardial injury following CME.
**Patients and Methods:** This study included 1198 colon cancer patients, 133 of whom experienced myocardial injury after surgery. Thirty-six distinct variables were gathered, encompassing patient demographics, medical history, preoperative examination characteristics, surgery type, and intraoperative details. Four machine learning algorithms, namely, extreme gradient boosting (XGBoost), random forest (RF), multilayer perceptron (MLP), and k-nearest neighbor algorithm (KNN), were employed to fabricate the model, and k-fold cross-validation, ROC curve, calibration curve, decision curve analysis (DCA), and external validation were employed to evaluate it.
**Results:** Out of the four predictive models employed, the XGBoost algorithm demonstrated the best performance. The ROC curve findings indicated that the XGBoost model exhibited remarkable predictive accuracy, with an area under the curve (AUC) value of 0.997 in the training set and 0.956 in the validation set. For internal validation, the k-fold cross-validation method was utilized, and the XGBoost model was shown to be steady. Furthermore, the calibration curves demonstrated the XGBoost model's high predictive capability. The DCA curve revealed higher benefit rates for patients who underwent interventional treatment under the XGBoost model. The AUC value for the external validation set was 0.74, which indicated that the XGBoost prediction model possessed good extrapolative capacity.
**Conclusion:** The myocardial injury prediction model for patients undergoing CME that was developed using the XGBoost machine learning algorithm in this study demonstrates both high predictive accuracy and clinical utility.
**Keywords:** colonic neoplasms, myocardial injury, surgery, prognosis, risk factor, machine learning

## Introduction

Colorectal cancer, a malignant tumor of the gastrointestinal tract, has a dismal prognosis. Lifestyle and dietary changes have contributed to a gradual increase in the incidence of colorectal cancer. Recent epidemiological studies have demonstrated that colorectal cancer is the third most common cancer worldwide, following lung and breast cancer.[1] Surgical intervention remains the mainstay of treatment for colorectal cancer, and Hohenberger proposed complete mesocolic excision (CME) as an effective procedure for reducing patient mortality.[2,3] This technique entails complete resection of the tumor and colonic mesentery, removal of surrounding lymph nodes, and full blockage of the main colonic vessels, which aligns with the principles of radical colorectal cancer surgery. Although clinicians are benefiting from radical colorectal cancer surgery, they are also observing that certain patients develop ischemia-reperfusion injury and hemodynamic changes after the procedure, leading to an elevated risk of myocardial injury in the postoperative period.

As the aging population undergoing surgery continues to grow, the incidence of such complications is expected to increase.[4] Multiple studies have indicated that perioperative cardiovascular events have become a leading cause of postoperative mortality, with approximately 1–2% of patients succumbing within 30 days after noncardiac surgery, which presents a challenging public health issue.[5–7]

A multicenter prospective study called the VISION study[6] was initiated globally and enrolled 40,000 middle-aged patients. The study highlighted the significance of elevated postoperative troponin levels in predicting myocardial injury and death in patients after noncardiac surgery. This finding further emphasized the importance of postoperative troponin testing for patients. However, universal troponin testing in all patients would impose a substantial burden on healthcare systems worldwide. As a result, the objective of this investigation was to identify patients with a heightened susceptibility to myocardial injury for subsequent testing, including troponin assessment. Conversely, patients with a low probability of myocardial injury are less likely to undergo troponin testing, thereby minimizing the financial burden on the hospital and alleviating the distress for the patient's family.

Surgeons often rely on their clinical experience for preoperative risk assessment of surgical patients at high risk for myocardial injury. However, the temporal nature and subjectivity of this approach often lead to inaccurate predictions. A decade ago, some researchers used relatively accurate parametric regression methods to predict disease, but the intricate relationships between clinical characteristic variables and the occurrence of postoperative complications render regression models inadequate for disease prediction. Moreover, regression models are limited by wide variation in predicted results when the number of cases of the variables in the model is small. Recently, with the rapid development of artificial intelligence in the medical field, machine learning has emerged as a widely used tool for clinical prediction and other applications.[8,9] In this study, we analyzed the clinical information of colon cancer patients and applied machine learning algorithms to develop a prediction model for myocardial injury after CME. Our aim is to identify high-risk patients and assist clinicians in making timely and accurate individualized treatment plans for patients.

## Materials and Methods

### Study Subjects

For this study, clinical data from the databases of Wuxi People's Hospital and Wuxi Second People's Hospital were utilized. The case inclusion criteria were as follows: (1) patients who underwent open CME or laparoscopic-assisted CME procedures, (2) cases in which the surgical team was composed of senior surgeons with the capability of performing CME independently, (3) patients with a confirmed postoperative pathological diagnosis of colon cancer, and (4) patients with a confirmed postoperative diagnosis of myocardial injury. Exclusion criteria for this study were as follows: (1) patients concomitant with other malignancies; (2) patients exhibiting distant metastasis of colon cancer through pathological examination or imaging; (3) patients diagnosed with acute heart failure, atrial fibrillation, and other serious cardiac dysfunctions prior to surgery; (4) patients with a history of open-heart surgery; and (5) patients with absent cases, clinical data, or visits. All patients who participated in this study were followed up for at least 3 years after the surgery (See Supplementary Material). The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of Wuxi People's Hospital and Wuxi Second People's Hospital with approval number KY22085.

### Study Design and Data Collection

The clinical information of patients diagnosed with colon cancer between 2010 and 2020 was procured from the databases of Wuxi People's Hospital and Wuxi Second People's Hospital. The data included 36 preoperative variables (within 24 h before the day of surgery), intraoperative variables, and postoperative variables (occurring 48 h after the initial surgery). Preoperative variables collected included patient demographic characteristics (gender, age, history of smoking, alcohol abuse, and body mass index), basic clinical characteristics (American Society of Anesthesiologists score, nutrition risk screening 2002 score, history of surgery, adjuvant chemotherapy history, and adjuvant radiotherapy history), basic medical history (anemia, tachycardia, chronic kidney disease, diabetes mellitus, hypertension, chronic obstructive pulmonary disease, and hyperlipidemia), laboratory tests (albumin, carcinoembryonic antigen and

carbohydrate antigen 19-9), and tumor characteristics (T-stage, N-stage, tumor size, and tumor number). The intraoperative variables collected included the type of surgery, number of lymph nodes dissected, duration of surgery, intraoperative bleeding, intraoperative blood transfusion, intraoperative percutaneous arterial oxygen saturation status, whether the patient experienced intraoperative tachycardia, and whether the surgery was an emergency procedure. Postoperative variables collected included laboratory test indices (neutrophil to lymphocyte ratio, procalcitonin, C-reactive protein, and serum amyloid A). The outcome variable of the study was postoperative myocardial injury.

## Diagnosis of Myocardial Injury

The diagnostic criteria for myocardial injury entail a troponin I measurement >45 ng/L and the patient fulfilling one or more of the following conditions: (1) the appearance of ischemic signs or symptoms in the neck, chest, and extremities; (2) the display of changes that indicate myocardial ischemia, such as ST-segment elevation on the ECG; (3) the presentation of pathological Q waves in any two consecutive leads for ≥30 ms on the ECG; (4) the occurrence of changes in the left branch block on the ECG; and (5) the exclusion of other underlying diseases that may cause myocardial injury, such as sepsis or pulmonary embolism.[10,11]

## Development and Evaluation of Predictive Models

In this study, SPSS and R software were employed for statistical analysis. To construct and evaluate the clinical prediction models, the following procedures were carried out: (1) Univariate and multivariate regression analyses were conducted. The chi-square test was utilized to compare the differences between two groups for categorical variables, while a *t*-test was applied for continuous variables that followed a normal distribution. The rank sum test was employed for continuous variables that did not conform to a normal distribution. The threshold for statistical significance was set at p<0.05. Logistic regression analysis was then performed on the significant variables from the univariate analysis to determine their independent influences on postoperative myocardial injury. Four models, namely, extreme gradient boosting (XGBoost), random forest (RF), multilayer perceptron (MLP), and k-nearest neighbor algorithm (KNN), were used to score the importance of each factor and rank them according to the weight of their influence. Only the variables that ranked in the top fifteen in all four model rankings and were significant in both univariate and multivariate analyses were selected. (2) Evaluate and build prediction models. Colon cancer patients diagnosed at Wuxi People's Hospital between January 2010 and January 2020 were chosen as the internal validation set, while those from Wuxi Second People's Hospital during the same period served as the external validation set. The internal validation set was randomly split into a training set (70%) and a test set (30%). After filtering out the clinical variables, four machine learning algorithms, namely, MLP, RF, XGBoost, and KNN, were applied. The performance of the models was evaluated using three criteria: discrimination, calibration, and clinical utility, and the best model was selected for further analysis. The model's predictive ability was determined using a receiver operating characteristic (ROC) curve to calculate the area under the curve (AUC) value. The calibration curve was plotted to assess the agreement between the predicted and actual results, while decision curve analysis (DCA) was performed to determine the benefit of interventional treatment for patients. Internal validation was performed using a k-fold cross-validation methodology. (3) External validation of the best model. The generalizability and predictive efficiency of the model was assessed by applying it to an external validation set and plotting ROC curves. (4) Model interpretation. SHAP (Shapley Additive Explanation) is a model visualization method used to explain individual predictions made by machine learning models. It provides a framework to calculate the contribution of each feature to the prediction made by the model for each instance in the dataset. The SHAP summary plot ranks the importance of the features based on their contribution to the model's predictions. It displays the SHAP values for each feature, with a bar for each feature indicating the feature's impact on the model's output. The longer the bar, the more important the feature is in determining the model's prediction. The SHAP force plot is used to explain the prediction of an individual sample. It displays the contribution of each feature to the prediction of the selected instance, with the sum of the contributions equaling the model's prediction for that instance. The force plot can help to understand why the model made a certain prediction for a specific sample and to identify which features were most influential in that prediction.

# Results

## Clinical Information of the Patients

A total of 1198 patients who had been diagnosed with colon cancer were included in this study, including 133 (11.1%) patients who suffered from postoperative myocardial injury (Figure 1).

## Screening for Risk Factors for Postoperative Myocardial Injury

The findings from both univariate and multivariate analyses demonstrated that age, body mass index (BMI), history of alcohol intake, smoking history, hyperlipidemia, previous episodes of tachycardia, diabetes mellitus history, chronic obstructive pulmonary disease (COPD) history, duration of surgery, intraoperative bleeding, intraoperative percutaneous arterial oxygen saturation (SPO$^2$) status, intraoperative blood transfusion, intraoperative tachycardia, tumor T-stage, tumor lymph node metastasis, tumor size, postoperative C-reactive protein (CRP) level, and postoperative neutrophil-to-lymphocyte ratio (NLR) were significant independent factors ($P<0.05$) influencing the outcome (Table 1). The risk factors affecting postoperative myocardial injury were identified by the XGBoost, RF, MLP, and KNN models, which included advanced age, preoperative and intraoperative tachycardia, BMI, smoking history, NLR, CRP, intraoperative blood transfusion, intraoperative SPO$^2$, operative time, and intraoperative bleeding (Figure 2A–D). A comprehensive analysis of these variables led to the selection of advanced age, preoperative and intraoperative tachycardia, BMI ≥25 kg/m$^2$, history of smoking, NLR ≥3, CRP ≥10 mg/l, intraoperative blood transfusion, intraoperative SPO$^2$ <90%, operative time ≥270 min, and intraoperative bleeding ≥100 ml as predictors for the model.
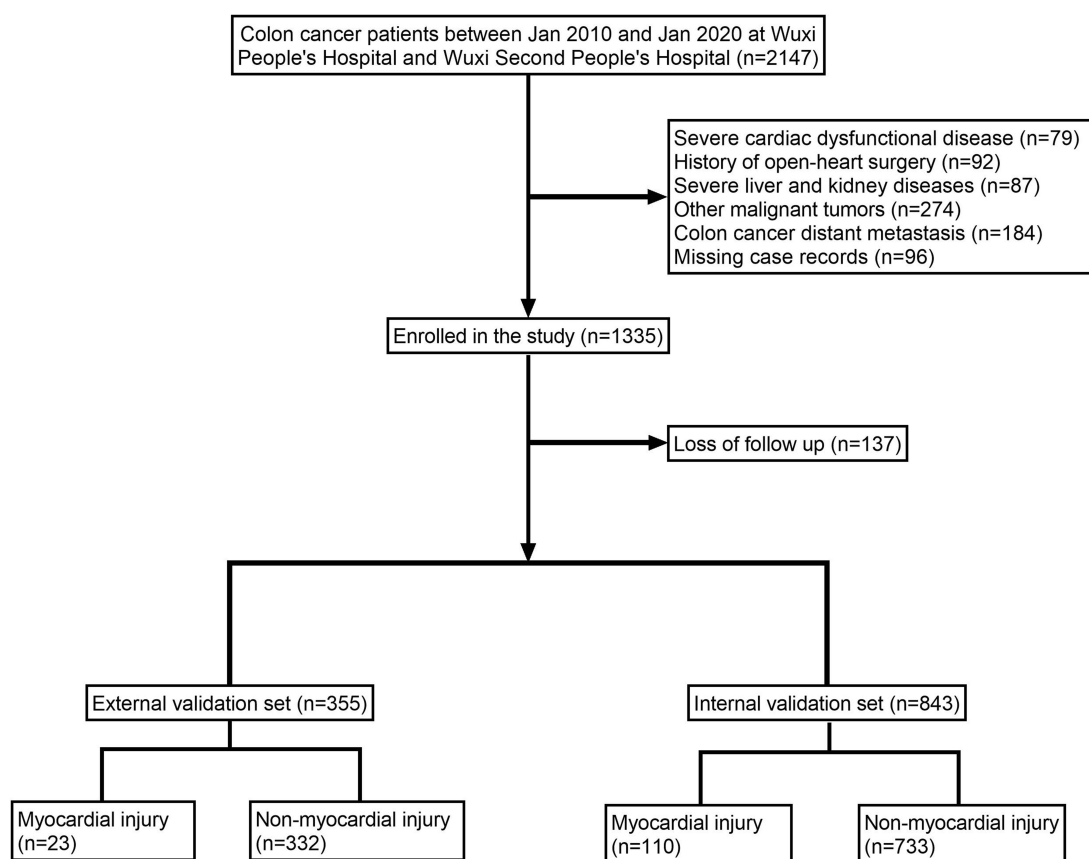


**Figure 1** Flow diagram of patients included in the study.

**Table 1** Univariate and Multivariate Analyses of Variables Related to Postoperative Myocardial Injury

| Variables | | Univariate Analysis | | | Multivariate Analysis | | |
|---|---|---|---|---|---|---|---|
| | | OR | 95%CI | P-value | OR | 95%CI | P-value |
| Sex | Female | Reference | | | | | |
| | Male | 1.123 | [0.748, 1.687] | 0.575 | | | |
| Age | <65 | Reference | | | Reference | | |
| | ≥65 | 10.832 | [6.811, 17.228] | <0.001 | 4.655 | [1.711, 13.402] | 0.003 |
| BMI | <25 kg/m² | Reference | | | Reference | | |
| | ≥25 kg/m² | 4.151 | [2.744, 6.280] | <0.001 | 4.268 | [1.698, 11.38] | 0.003 |
| ASA | <3 | Reference | | | | | |
| | ≥3 | 1.046 | [0.700, 1.563] | 0.828 | | | |
| ALB | ≥30g/L | Reference | | | Reference | | |
| | <30g/L | 3.3 | [2.146, 5.075] | <0.001 | 2.384 | [0.901, 6.503] | 0.082 |
| NRS2002 score | <3 | Reference | | | | | |
| | ≥3 | 0.771 | [0.483, 1.231] | 0.276 | | | |
| Drinking history | No | Reference | | | Reference | | |
| | Yes | 4.163 | [2.740, 6.326] | <0.001 | 3.144 | [1.288, 8.031] | 0.013 |
| Smoking history | No | Reference | | | Reference | | |
| | Yes | 9.308 | [6.018, 14.398] | <0.001 | 10.163 | [2.945, 41.443] | 0.001 |
| Surgical history | No | Reference | | | | | |
| | Yes | 1.257 | [0.837, 1.887] | 0.271 | | | |
| Anemia | No | Reference | | | | | |
| | Yes | 1.191 | [0.767, 1.848] | 0.436 | | | |
| Hyperlipidemia | No | Reference | | | Reference | | |
| | Yes | 2.787 | [1.855, 4.188] | <0.001 | 2.872 | [1.161, 7.473] | 0.025 |
| Hypertension | No | Reference | | | Reference | | |
| | Yes | 6.082 | [3.932, 9.407] | <0.001 | 2.03 | [0.765, 5.508] | 0.157 |
| Diabetes | No | Reference | | | Reference | | |
| | Yes | 6.327 | [4.135, 9.682] | <0.001 | 18.837 | [7.069, 57.242] | <0.001 |
| Preoperative tachycardia | No | Reference | | | Reference | | |
| | Yes | 4.373 | [2.887, 6.625] | <0.001 | 9.24 | [3.536, 27.134] | <0.001 |
| COPD | No | Reference | | | Reference | | |
| | Yes | 6.198 | [4.031, 9.529] | <0.001 | 4.861 | [1.334, 18.447] | 0.017 |
| CKD | No | Reference | | | | | |
| | Yes | 1.33 | [0.800, 2.209] | 0.271 | | | |
| Adjuvant Radiotherapy | No | Reference | | | | | |
| | Yes | 1.491 | [0.923, 2.409] | 0.103 | | | |
| Adjuvant Chemotherapy | No | Reference | | | | | |
| | Yes | 1.008 | [0.648, 1.568] | 0.971 | | | |
| Surgical procedure | Laparoscopic surgery | Reference | | | | | |
| | Open surgery | 0.688 | [0.460, 1.031] | 0.07 | | | |
| Emergency surgery | No | Reference | | | | | |
| | Yes | 1.067 | [0.694, 1.642] | 0.767 | | | |
| Lymph node dissection | <12 | Reference | | | | | |
| | ≥12 | 0.805 | [0.514, 1.261] | 0.344 | | | |
| Surgery time | <270 min | Reference | | | Reference | | |
| | ≥270 min | 6.901 | [4.492, 10.602] | <0.001 | 4.202 | [1.683, 11.11] | 0.003 |
| Intraoperative bleeding | <100 ml | Reference | | | Reference | | |
| | ≥100 ml | 4.626 | [3.053, 7.009] | <0.001 | 4.288 | [1.639, 11.914] | 0.004 |
| SpO² | ≥90% | Reference | | | Reference | | |
| | <90% | 4.081 | [2.695, 6.180] | <0.001 | 5.438 | [2.15, 14.995] | 0.001 |
| Blood transfusion | No | Reference | | | Reference | | |
| | Yes | 3.343 | [2.073, 5.391] | <0.001 | 5.491 | [1.758, 18.277] | 0.004 |
| Intraoperative tachycardia | No | Reference | | | Reference | | |
| | Yes | 4.507 | [2.960, 6.864] | <0.001 | 4.838 | [1.914, 13.176] | 0.001 |
| T-stage | T1~T2 | Reference | | | Reference | | |
| | T3~T4 | 4.343 | [2.843, 6.636] | <0.001 | 3.706 | [1.44, 9.97] | 0.007 |

*(Continued)*

**Table 1** (Continued).

| Variables | | Univariate Analysis | | | Multivariate Analysis | | |
|---|---|---|---|---|---|---|---|
| | | OR | 95%CI | P-value | OR | 95%CI | P-value |
| N-stage | N0 | Reference | | | Reference | | |
| | N1~N2 | 4.082 | [2.688, 6.198] | <0.001 | 3.097 | [1.126, 8.992] | 0.031 |
| Tumor number | <2 | Reference | | | | | |
| | ≥2 | 1.017 | [0.611, 1.695] | 0.947 | | | |
| Tumor size | <5 cm | Reference | | | Reference | | |
| | ≥5 cm | 3.062 | [1.926, 4.868] | <0.001 | 3.658 | [1.177, 11.87] | 0.027 |
| CEA level | <5 ng/ml | Reference | | | | | |
| | ≥5 ng/ml | 1.11 | [0.677, 1.821] | 0.679 | | | |
| CA19-9 level | <37 U/mL | Reference | | | | | |
| | ≥37 U/mL | 0.77 | [0.464, 1.276] | 0.31 | | | |
| PCT level | <0.05 ng/ml | Reference | | | | | |
| | ≥0.05 ng/ml | 1.169 | [0.771, 1.772] | 0.461 | | | |
| CRP level | <10 mg/l | Reference | | | Reference | | |
| | ≥10 mg/l | 2.236 | [1.485, 3.366] | <0.001 | 2.891 | [1.115, 7.863] | 0.032 |
| SAA level | <10 mg/l | Reference | | | | | |
| | ≥10 mg/l | 0.851 | [0.570, 1.272] | 0.432 | | | |
| NLR | <3 | Reference | | | Reference | | |
| | ≥3 | 3.219 | [2.090, 4.959] | <0.001 | 4.646 | [1.793, 12.782] | 0.002 |

**Abbreviations**: OR, odds ratio; CI, confidence interval; BMI, body mass index; ASA, The American Society of Anesthesiologists; ALB, albumin; CA19-9, carbohydrate antigen 19-9; PCT, procalcitonin; CRP, C-reactive protein; SAA, serum amyloid A; NRS2002, nutrition risk screening 2002; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; $SPO_2$, percutaneous arterial oxygen saturation; NLR, neutrophil to lymphocyte ratio.

## Model Building and Evaluation

The training set's ROC curve results demonstrated XGBoost's exceptional performance, with an AUC value as high as 0.997. In the validation set, XGBoost outperformed the other three models, with an AUC value of 0.956 (Table 2). The calibration curves of all four models showed high consistency between the predicted and actual results, aligning with the ideal curves. Furthermore, the DCA curves indicated that all models exhibited a net clinical benefit in comparison to either a full or no treatment plan (Figure 3A–D). The generalization ability of the four models was assessed using k-fold cross-validation. In this investigation, the test set comprised N=253 cases (30.01%), while the remaining samples were utilized for 5-fold cross-validation. The XGBoost algorithm displayed a superior performance with an AUC value of 0.9478±0.0056 in the validation set and an AUC value of 0.9520 in the test set, alongside an accuracy of 0.9051 (Figure 4A–C). The RF algorithm demonstrated an AUC value of 0.7835±0.0659 in the validation set and an AUC value of 0.8464 in the test set, with an accuracy of 0.8577. The MLP algorithm exhibited an AUC value of 0.8910±0.0562 in the validation set and an AUC value of 0.9360 in the test set, with an accuracy of 0.8656. Finally, the KNN algorithm showed an AUC value of 0.8812±0.0476 in the validation set and an AUC value of 0.8870 in the test set, with an accuracy of 0.9051. Upon comprehensive comparison, the XGBoost algorithm was chosen for model construction in this study.

## Model External Validation

The external validation set demonstrated an AUC value of 0.74, indicating that the disease prediction model exhibited a high level of accuracy. (Figure 4D).

## Model Explanation

The SHAP summary plot depicted the importance of risk factors for myocardial injury after CME, with advanced age, operative time ≥270 min, intraoperative tachycardia, history of smoking, intraoperative $SPO_2$ <90%, BMI ≥25 kg/m$^2$, intraoperative bleeding ≥100 ml, preoperative tachycardia, CRP ≥10 mg/l, NLR ≥3, and intraoperative blood transfusion ranking as the top 11 factors (Figure 5). The SHAP force plot illustrates the predictive analysis of the model for two patients with myocardial injury. The model predicted a probability of myocardial injury of 0.412 for patient I, and the
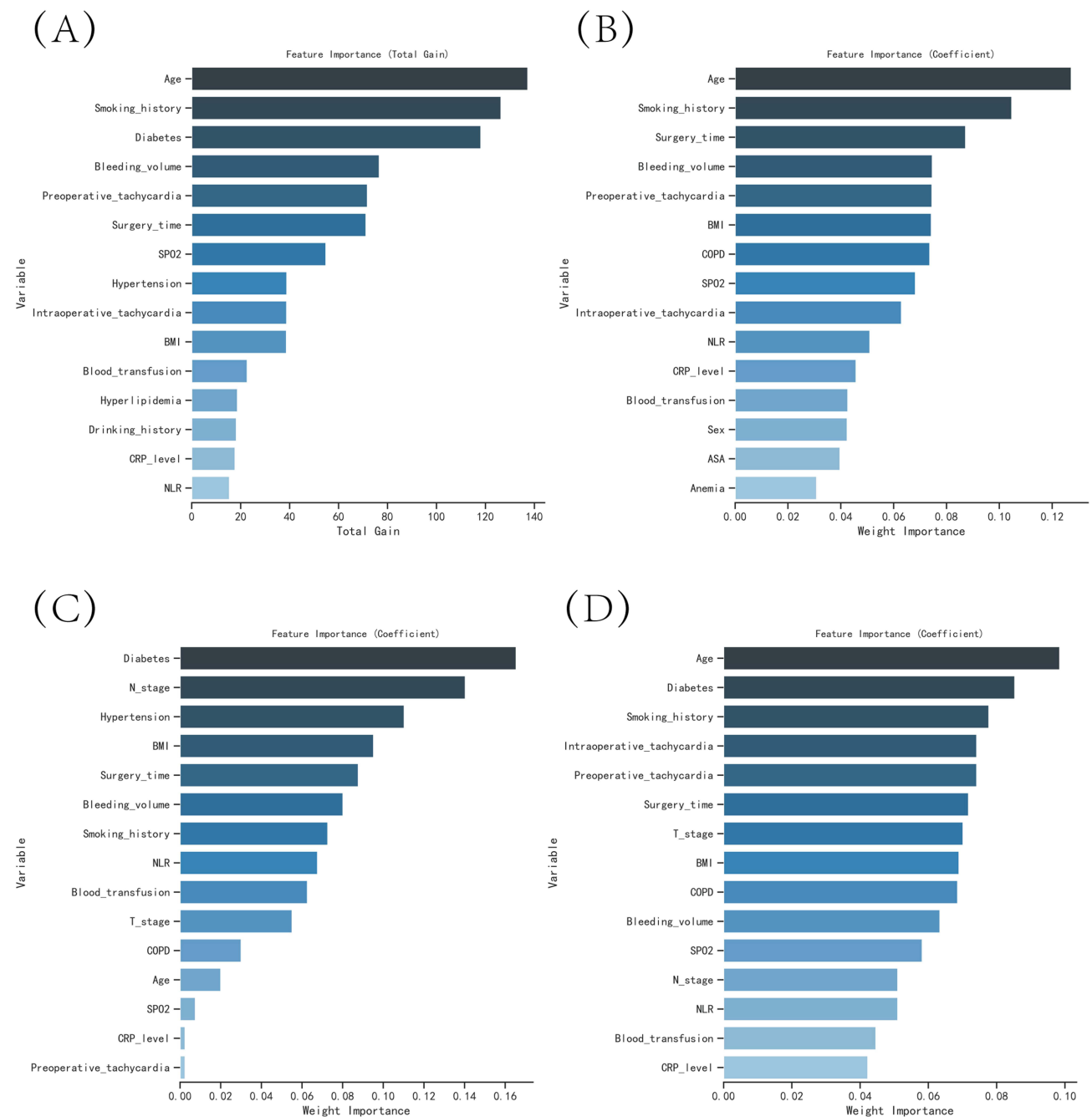
**Figure 2** The variable influence factor ranking plots of the four models. (**A**) Variable importance ranking diagram of the XGBoost model. (**B**) Variable importance ranking diagram of the RF model. (**C**) Variable importance ranking diagram of the MLP model. (**D**) Variable importance ranking diagram of the KNN model.

factors that increased its likelihood were preoperative tachycardia, intraoperative tachycardia, NLR $\geq$ 3, and operative time $\geq$ 270 min. The model predicted a probability of myocardial injury of 0.854 for patient II, and the factors that increased its likelihood were intraoperative bleeding $\geq$ 100 ml, CRP $\geq$ 10 mg/l, intraoperative tachycardia, operative time $\geq$ 270 min, and advanced age (Figure 6A and B).

## Discussion

This study assesses the risk prediction models derived from four machine learning algorithms. Our findings indicate that the XGBoost algorithm demonstrates high accuracy, efficiency, flexibility, and universal applicability. In comparison to

**Table 2** Evaluation of the Four Models

|  |  | AUC (95%CI) | Accuracy (95%CI) | Sensitivity (95%CI) | Specificity (95%CI) |
|---|---|---|---|---|---|
| KNN | Training set | 0.981 (0.972–0.990) | 0.951(0.944–0.959) | 1.000(1.000–1.000) | 0.900(0.895–0.905) |
|  | Validation set | 0.863 (0.740–0.979) | 0.927(0.912–0.943) | 0.763(0.722–0.803) | 0.931(0.878–0.984) |
| XGBoost | Training set | 0.997 (0.993–1.000) | 0.974(0.970–0.978) | 0.977(0.964–0.990) | 0.976(0.971–0.981) |
|  | Validation set | 0.956 (0.920–0.991) | 0.910(0.884–0.936) | 0.973(0.941–1.005) | 0.883(0.829–0.937) |
| RF | Training set | 0.816 (0.761–0.871) | 0.826(0.781–0.871) | 0.837(0.778–0.896) | 0.685(0.596–0.774) |
|  | Validation set | 0.814 (0.697–0.931) | 0.837(0.788–0.886) | 0.772(0.679–0.865) | 0.788(0.712–0.865) |
| MLP | Training set | 0.964 (0.945–0.984) | 0.902(0.865–0.939) | 0.907(0.878–0.936) | 0.902(0.859–0.946) |
|  | Validation set | 0.940 (0.882–0.997) | 0.895(0.872–0.918) | 0.942(0.864–1.019) | 0.862(0.815–0.909) |

**Abbreviations**: AUC, area under the curve; RF, random forest; XGBoost, extreme gradient boosting; MLP, multilayer perceptron; KNN, k-nearest neighbor algorithm; CI, confidence interval.

the RF algorithm, the XGBoost algorithm effectively addresses the regularization problem and mitigates model overfitting.[12] While the MLP algorithm and KNN algorithm are more precise and proficient at avoiding overfitting, they exhibit less stability when tackling multi feature and large sample problems.[13] The XGBoost algorithm is better suited for multidimensional studies and substantially reduces computational effort and training time. In contrast to the SVM algorithm and KNN algorithm, the XGBoost algorithm possesses several distinct advantages. Accordingly, after scrutinizing all four machine learning algorithms, the XGBoost algorithm was deemed optimal for constructing a model aimed at predicting the incidence of myocardial injury after CME in this investigation.

A number of studies[14,15] have partially validated the effectiveness of machine learning algorithms in clinical diagnosis and prognosis. These techniques have been shown to predict adverse outcomes in disease progression with greater accuracy than traditional diagnostic methods. In the present study, machine learning algorithms played a crucial role in developing a predictive model. This model can reduce the number of nonessential tests that patients undergo, lessen the financial burden on their families, and minimize the side effects of diagnostic instruments on patients. Most importantly, this model can enable clinical decision makers to accurately identify high-risk patients and initiate timely interventional treatments to improve patient prognosis. In this study, SHAP analysis was employed to identify the risk factors for myocardial injury after CME, including advanced age, smoking history, preoperative tachycardia, intraoperative tachycardia, CRP levels $\geq 10$ mg/L, an NLR $\geq 3$, BMI $\geq 25$ kg/m$^2$, SPO$^2$ levels <90%, intraoperative blood transfusion, long operative time, and high intraoperative bleeding.

As patients age, they become more susceptible to developing organic lesions in the heart. The myocardium of elderly patients undergoes varying degrees of pathological changes, such as hypertrophy and fibrosis, which ultimately affect cardiac pumping function. Moreover, elderly patients are particularly vulnerable to myocardial injury due to changes in the sympathetic and endocrine systems caused by major surgical procedures such as CME. This results in an increased volume and pressure load on the heart. Patients with preoperative tachycardia and those who experience intraoperative tachycardia are also at a higher risk of developing myocardial injury in the postoperative period, according to the findings of this study. It is widely believed that surgery, as a specific form of trauma, can worsen preexisting complications in patients. Patients with tachycardia are at a higher risk of myocardial injury, and those with atypical clinical presentations may develop cardiovascular diseases such as acute heart failure in response to triggers such as surgery. Furthermore, the surgical procedure can cause an increase in pressure in the patient's left atrium, significantly increasing the risk of postoperative myocardial injury. Studies conducted by Philippe et al have demonstrated a robust correlation between the patient's heart rate and postoperative heart failure. Specifically, for every 10 beats/min increase in heart rate, the risk of postoperative heart failure increases by 10%.[16,17] This further supports the notion that rapid heart rate can be considered a new risk factor in predicting postoperative myocardial injury.

This study utilized two samples to explore models for predicting postoperative myocardial injury in patients. In both samples, longer operative time and higher intraoperative bleeding were found to be significant risk factors. This may be attributed to the large fluid loss and prolonged exposure to anesthetic drugs, which increase the workload of myocardial cells and decrease oxygen supply, ultimately leading to an increased risk of myocardial injury. Prolonged surgery also
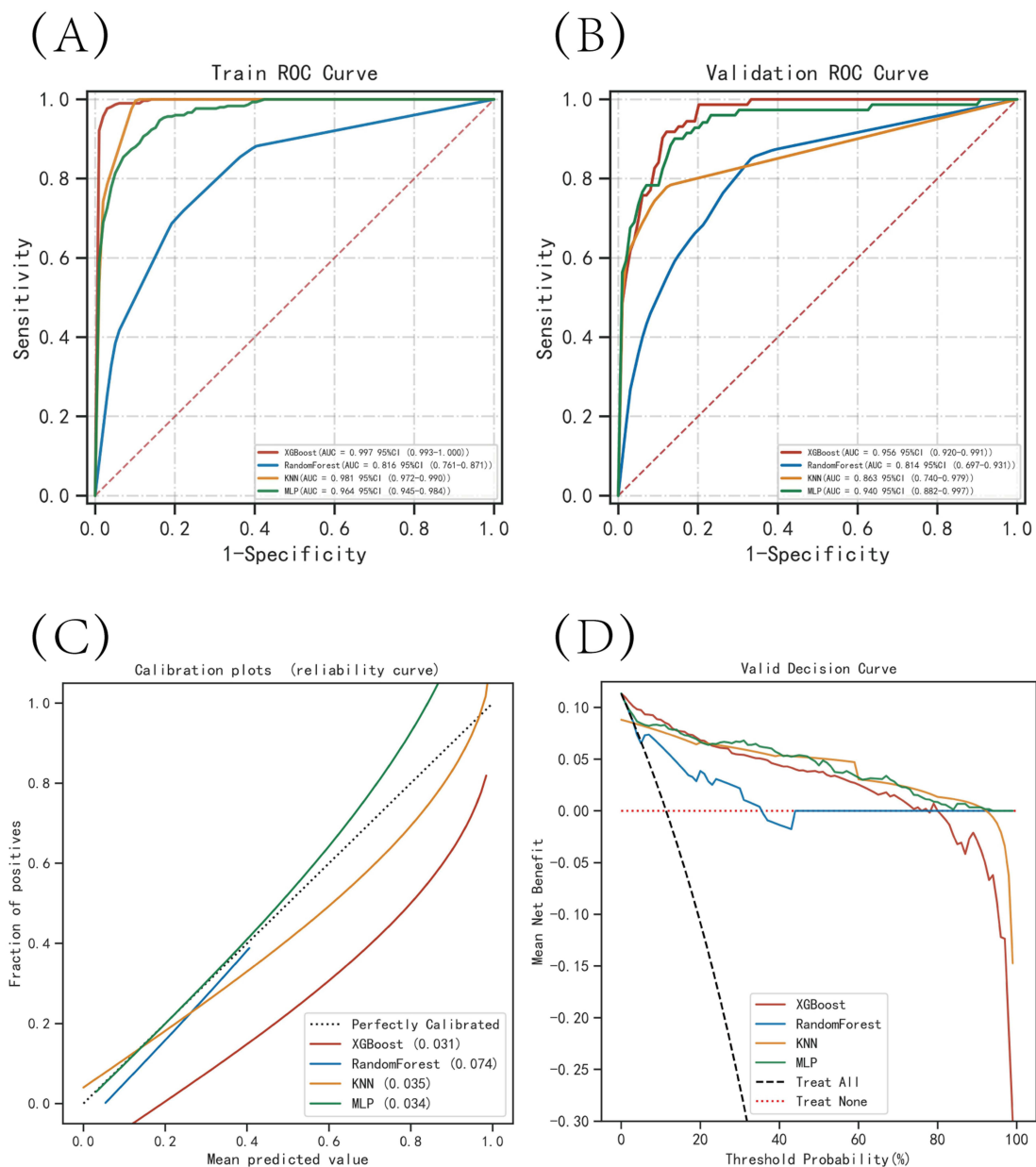
**Figure 3** Evaluation of the four models for predicting myocardial injury. (**A**) ROC curves for the training set of the four models. (**B**) ROC curves for the validation set of the four models. (**C**) Calibration plots of the four models. In a calibration plot, the 45° dotted line represents perfect calibration, where the observed and predicted values are perfectly matched. The distance between the observed and predicted values is shown by the curves, and a closer distance indicates greater accuracy of the model. (**D**) DCA curves of the four models. In a decision curve analysis, the x-axis represents the threshold probability of the outcome, while the y-axis represents the net benefit of making a decision based on the model's prediction. The "All" curve represents the net benefit of treating all patients, regardless of the model's prediction, while the "None" curve represents the net benefit of not treating any patients.

triggers the hypothalamic–pituitary–adrenal (HPA) axis and sympathetic nervous system, resulting in the release of cortisol and catecholamines, which affect heart rate, blood pressure, and myocardial contractility. Although these hormones may help the body cope with intraoperative injury in the short term, long-term activation may cause potential damage to cardiac myocytes. Another important mechanism of myocardial injury is the immune and inflammatory response triggered by surgery. Inflammatory factors play a role in tissue healing and repair but may also have damaging effects.[18,19] Studies have suggested that certain inflammatory factors, such as tumor necrosis factor-α and interleukin-1, can lead to oxidative stress and mitochondrial dysfunction in cardiac myocytes, disrupting normal cellular function. These findings support the conclusions drawn from the present study.
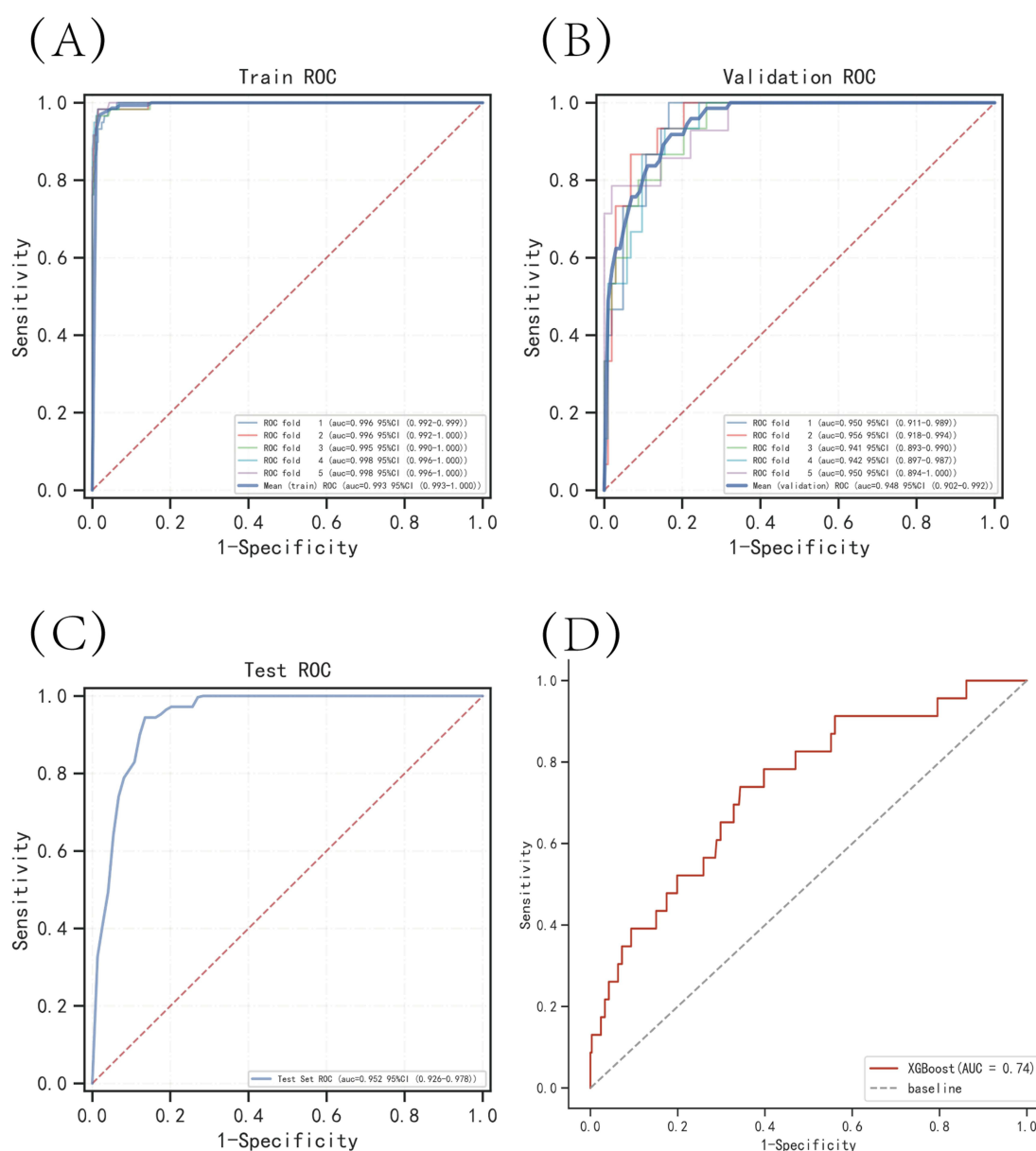
**Figure 4** Internal validation of the XGBoost model. (**A**) ROC curve of the XGBoost model for the training set. (**B**) ROC curve of the XGBoost model for the validation set. (**C**) ROC curve of the XGBoost model for the test set. (**D**) External validation of the XGBoost model.

This investigation revealed that obese individuals were more inclined to suffer from myocardial injury subsequent to surgery. These patients are more often combined with chronic diseases, such as hypertension and diabetes, which increase the risk of myocardial injury.[20–22] In addition, obese patients are more prone to perivascular adipose tissue, resulting in a greater risk of surgical injury and presenting difficulties with respect to local exposure of the surgical field during open CME or laparoscopic-assisted CME procedures. These factors exacerbate the complexity of surgery, necessitating greater surgical expertise and predisposing the patient to a greater likelihood of intraoperative bleeding, which increases the hemodynamic load on the patient. Zhou et al discovered that high BMI is an autonomous risk factor for cardiovascular disease.[23] Moreover, this study also determined that chronic smoking is a significant risk factor for postoperative myocardial injury. Patients who smoke experience further elevations in cardiopulmonary burden and are at risk of complications, such as pulmonary stasis and compensatory hypertrophy of the myocardium, which impairs cardiac function. Furthermore, patients with altered airway structures and diminished immune defenses are susceptible to pulmonary infections and more vulnerable to cardiovascular complications.[24] Consequently, healthcare professionals are encouraged to offer active health education to
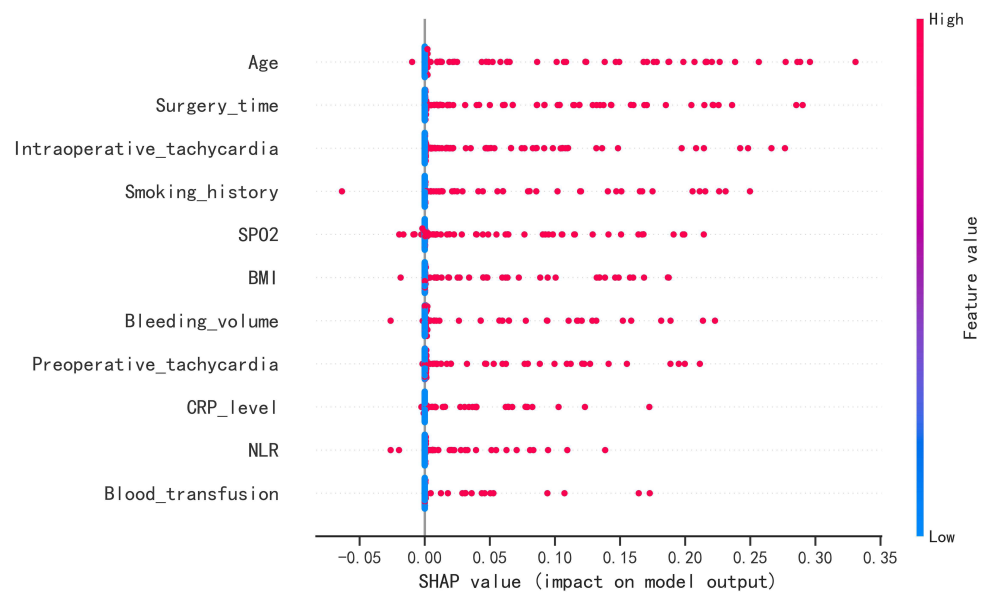
**Figure 5** SHAP summary plot. Risk factors are arranged along the y-axis based on their importance, which is given by the mean of their absolute Shapley values. The higher the risk factor is positioned in the plot, the more important it is for the model.
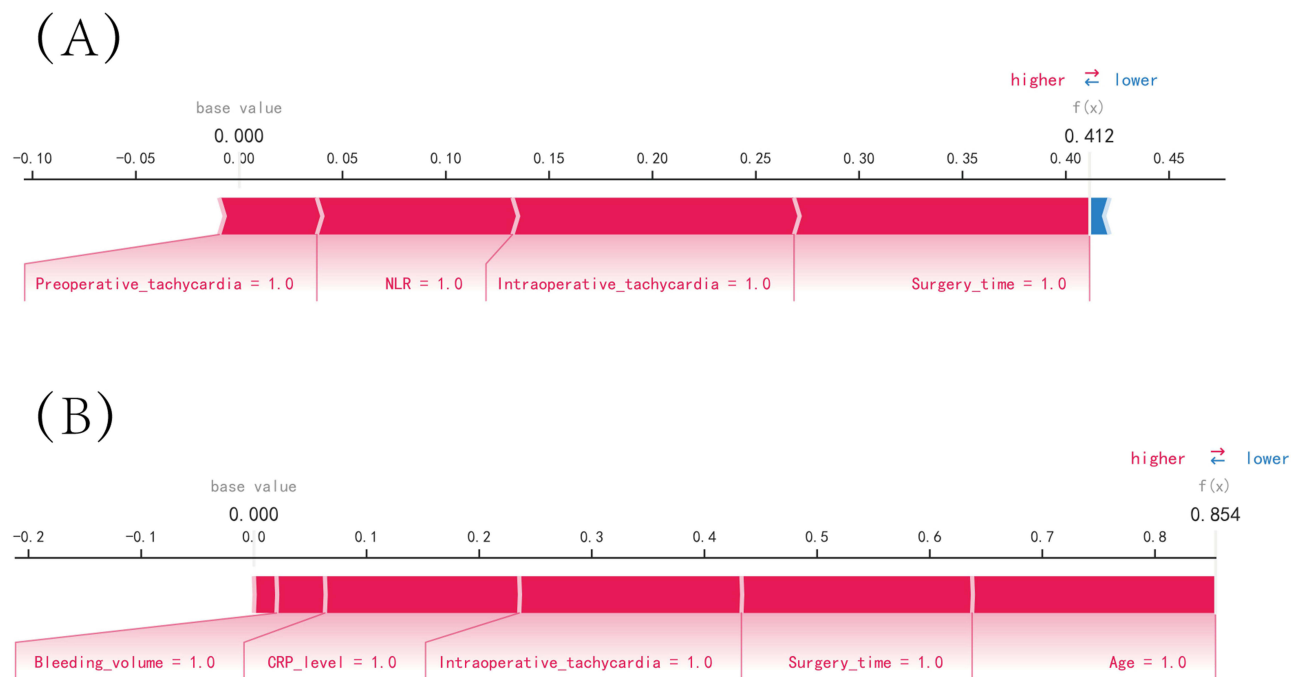


**Figure 6** SHAP force plot. SHAP analysis is a method used to explain the output of a machine learning model by attributing the contribution of each feature to the final prediction. The SHAP values represent the contribution of each feature to the difference between the actual prediction and the baseline prediction. In the horizontal line plot, each feature is represented by a bar with a corresponding SHAP value. The position of the bar along the horizontal line indicates the magnitude and direction of the impact of the feature on the prediction. The features are sorted by their absolute SHAP value, with the most important features on the left and the least important on the right. The color of the bar represents the direction of the feature's impact. Blue bars indicate features that have a negative effect on the disease prediction, meaning that an increase in the feature value decreases the likelihood of disease. Red bars indicate features that have a positive effect on disease prediction, meaning that an increase in the feature value increases the likelihood of disease. (**A**) Predictive Analysis of Patient I. (**B**) Predictive Analysis of Patient II.

patients, advise them to quit smoking two to three weeks before the surgery, and encourage appropriate physical activity to enhance cardiopulmonary function. Patients should be closely observed for signs of pulmonary distress and laboratory indicators, and prophylactic antimicrobial drugs should be administered to deter pulmonary infections.

In recent times, certain researchers have expressed optimism that alternative methods to troponin testing could potentially curtail healthcare expenditures. However, this approach has led to a proportionate escalation in misdiagnoses and heightened mortality risk among patients.[5–7] Consequently, medical practitioners have commenced the process of refining their diagnostic protocols by leveraging predictive models that factor in cardiovascular disease. However, these studies tend to focus solely on the general clinical characteristics of patients.[25,26] Notably, the development of post-operative myocardial injury can be influenced by the neurological and endocrine systems, as alterations in these systems may cause overexpression of proinflammatory cytokines and signaling molecules. Thus, the present study considered relevant inflammatory indicators to assess their impact on the development of myocardial injury. Our study demonstrated a higher incidence of myocardial injury in patients with elevated postoperative NLR and CRP. A meta-analysis of 14 studies also supported the strong association between inflammatory markers and the development of cardiovascular disease.[27] Additionally, Christian employed CRP to predict cardiovascular function in patients and provided precise individualized treatment plans. This further reinforces the findings of our current study.[28–30]

This study also incorporated some high-risk factors that may be overlooked by clinicians. The findings revealed that an $SPO^2$ level below 90% was a significant risk factor for postoperative myocardial injury. Although cardiomyocytes are less sensitive to hypoxia than brain cells, they are still more vulnerable to injury during intraoperative hypoxia. The severity and duration of myocardial injury may depend on the severity and duration of hypoxia, as suggested by previous studies.[31–33] When cardiomyocytes continue to pump blood at a high rate without adequate oxygen, lactic acid accumulation may result in acidosis, leading to cellular death. Furthermore, hypoxia can trigger biochemical responses such as mitochondrial dysfunction and the release of reactive oxygen species (ROS) and cytokines, which increase the risk of myocardial injury.[34] This study also revealed that patients who received blood transfusions during surgery were also at a higher risk of postoperative myocardial injury. A study by Ronald et al demonstrated that blood transfusions may trigger immunosuppression and increase the incidence of postoperative infection in oncology patients.[35] Additionally, blood transfusions may lead to the formation of reactive oxygen species, which can have a destructive effect on cell membranes and other cellular components.

## Limitations

The present investigation offers a thorough assessment of three facets pertaining to the model, namely discrimination, calibration, and clinical utility. This evaluation serves as an exemplar for a predictive model that prognosticates postoperative complications in other moderately-to-highly complex abdominal surgeries. However, there were some limitations. For instance, the study did not focus on biochemical indicators of patients, despite including several risk factors. Moreover, although the machine learning algorithms used had high accuracy, the resulting models were more complex and less interpretable. The entire computational and decision-making process of the model is performed in a black box, which is not as intuitive and clear as logistic regression models.[36–38] Additionally, the current study was a retrospective study, which has disadvantages such as selection bias and retrospective bias. Therefore, future studies should include multicenter prospective studies to enhance the reliability of the results.

## Conclusion

In this study, an XGBoost machine learning algorithm was used to construct a model for predicting the risk of myocardial injury after CME. The model has high prediction accuracy and clinical utility and can provide strong support for surgeons to diagnose patients in a timely manner. The results showed that myocardial injury remains one of the major problems faced by CME patients after surgery and is closely associated with advanced age, smoking history, preoperative tachycardia, intraoperative tachycardia, CRP ≥10 mg/l, NLR ≥3, BMI ≥25 kg/m², $SPO^2$ <90%, intraoperative blood transfusion, long operative time, and high intraoperative bleeding.

## Data Sharing Statement

The original data presented in the study are included in <u>Supplementary Material</u>, and further inquiries can be directed to the corresponding author (shenweijs@outlook.com).

## Ethics Statement

This study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of Wuxi People's Hospital, with approval number KY22085. The review committee waived the requirement for written informed consent because of the retrospective nature of the study. Prior to analysis, confidential patient information was deleted from the entire data set.

## Acknowledgments

The authors express gratitude to their affiliated institutions for their backing in the composition of this investigation.

## Disclosure

The authors report no conflicts of interest in this work.

## References

1. Mattiuzzi C, Sanchis-Gomar F, Lippi G. Concise update on colorectal cancer epidemiology. *Ann Transl Med*. 2019;7(21):609. doi:10.21037/atm.2019.07.91
2. Hohenberger W, Weber K, Matzel K, Papadopoulos T, Merkel S. Standardized surgery for colonic cancer: complete mesocolic excision and central ligation--technical notes and outcome. *Colorectal Dis*. 2009;11(4):354–64;discussion 64–5. doi:10.1111/j.1463-1318.2008.01735.x
3. Jayne DG, Thorpe HC, Copeland J, Quirke P, Brown JM, Guillou PJ. Five-year follow-up of the Medical Research Council CLASICC trial of laparoscopically assisted versus open surgery for colorectal cancer. *Br J Surg*. 2010;97(11):1638–1645. doi:10.1002/bjs.7160
4. van Waes JA, Nathoe HM, de Graaff JC, et al. Myocardial injury after noncardiac surgery and its association with short-term mortality. *Circulation*. 2013;127(23):2264–2271. doi:10.1161/CIRCULATIONAHA.113.002128
5. Devereaux PJ, Yang H, Yusuf S, et al. Effects of extended-release metoprolol succinate in patients undergoing non-cardiac surgery (POISE trial): a randomised controlled trial. *Lancet*. 2008;371(9627):1839–1847.
6. Devereaux PJ, Chan MT, Alonso-Coello P, et al. Association between postoperative troponin levels and 30-day mortality among patients undergoing noncardiac surgery. *JAMA*. 2012;307(21):2295–2304.
7. Fleisher LA, Eagle KA, Shaffer T, Anderson GF. Perioperative- and long-term mortality rates after major vascular surgery: the relationship to preoperative testing in the medicare population. *Anesth Analg*. 1999;89(4):849–855. doi:10.1213/00000539-199910000-00007
8. Croner RS, Ptok H, Merkel S, Hohenberger W. Implementing complete mesocolic excision for colon cancer - mission completed? *Innov Surg Sci*. 2018;3(1):17–29. doi:10.1515/iss-2017-0042
9. Katiyar A, Mohanty A, Hua J, et al. A Bayesian approach to determine the composition of heterogeneous cancer tissue. *BMC Bioinform*. 2018;19 (Suppl 3):90. doi:10.1186/s12859-018-2062-0
10. Zahid JA, Orhan A, Ekeloef S, Gögenur I. Myocardial injury after colorectal cancer surgery and postoperative 90-day mortality and morbidity: a retrospective cohort study. *Dis Colon Rectum*. 2021;64(12):1531–1541. doi:10.1097/DCR.0000000000002061
11. Devereaux PJ, Duceppe E, Guyatt G, et al. Dabigatran in patients with myocardial injury after non-cardiac surgery (MANAGE): an international, randomised, placebo-controlled trial. *Lancet*. 2018;391(10137):2325–2334. doi:10.1016/S0140-6736(18)30832-8
12. Wang L, Wang X, Chen A, Jin X, Che H. Prediction of type 2 diabetes risk and its effect evaluation based on the XGBoost model. *Healthcare*. 2020;8(3):154.
13. Hao PY, Chiang JH, Chen YD. Possibilistic classification by support vector networks. *Neural Netw*. 2022;149:40–56. doi:10.1016/j.neunet.2022.02.007
14. Karadaghy OA, Shew M, New J, Bur AM. Development and assessment of a machine learning model to help predict survival among patients with oral squamous cell carcinoma. *JAMA Otolaryngol Head Neck Surg*. 2019;145(12):1115–1120. doi:10.1001/jamaoto.2019.0981
15. Tseng YJ, Wang HY, Lin TW, Lu JJ, Hsieh CH, Liao CT. Development of a machine learning model for survival risk stratification of patients with advanced oral cancer. *JAMA Netw Open*. 2020;3(8):e2011768. doi:10.1001/jamanetworkopen.2020.11768
16. Steg PG, Dabbous OH, Feldman LJ, et al. Determinants and prognostic impact of heart failure complicating acute coronary syndromes: observations from the Global Registry of Acute Coronary Events (GRACE). *Circulation*. 2004;109(4):494–499. doi:10.1161/01.CIR.0000109691.16944.DA
17. O'Connor CM, Hathaway WR, Bates ER, et al. Clinical characteristics and long-term outcome of patients in whom congestive heart failure develops after thrombolytic therapy for acute myocardial infarction: development of a predictive model. *Am Heart J*. 1997;133(6):663–673. doi:10.1016/S0002-8703(97)70168-6
18. Guo X, Yin H, Li L, et al. Cardioprotective role of tumor necrosis factor receptor-associated factor 2 by suppressing apoptosis and necroptosis. *Circulation*. 2017;136(8):729–742. doi:10.1161/CIRCULATIONAHA.116.026240
19. Mihalovic M, Mikulenka P, Línková H, et al. Tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) in patients after acute stroke: relation to stroke severity, myocardial injury, and impact on prognosis. *J Clin Med*. 2022;11(9). doi:10.3390/jcm11092552
20. Stamler J. Epidemiologic findings on body mass and blood pressure in adults. *Ann Epidemiol*. 1991;1(4):347–362. doi:10.1016/1047-2797(91)90045-E

21. Chan JM, Rimm EB, Colditz GA, Stampfer MJ, Willett WC. Obesity, fat distribution, and weight gain as risk factors for clinical diabetes in men. *Diabetes Care*. 1994;17(9):961–969. doi:10.2337/diacare.17.9.961

22. Colditz GA, Willett WC, Rotnitzky A, Manson JE. Weight gain as a risk factor for clinical diabetes mellitus in women. *Ann Intern Med*. 1995;122 (7):481–486. doi:10.7326/0003-4819-122-7-199504010-00001

23. Zhou YT, Grayburn P, Karim A, et al. Lipotoxic heart disease in obese rats: implications for human obesity. *Proc Natl Acad Sci U S A*. 2000;97 (4):1784–1789. doi:10.1073/pnas.97.4.1784

24. Fuse K, Tanaka T, Shibasaki Y, et al. Marker chromosome is a strong poor prognosis factor after allogeneic HSCT for adverse-risk AML patients. *Eur J Haematol*. 2020;105(5):616–625. doi:10.1111/ejh.13495

25. Zheng L, Smith NJ, Teng BQ, Szabo A, Joyce DL. Predictive model for heart failure readmission using nationwide readmissions database. *Mayo Clin Proc Innov Qual Outcomes*. 2022;6(3):228–238. doi:10.1016/j.mayocpiqo.2022.04.002

26. de Oliveira Correia ET, Mechanick JI, Dos Santos Barbetta LM, Jorge AJL, Mesquita ET. Cardiometabolic-based chronic disease: adiposity and dysglycemia drivers of heart failure. *Heart Fail Rev*. 2023;28(1):47–61. doi:10.1007/s10741-022-10233-x

27. Zhang S, Diao J, Qi C, et al. Predictive value of neutrophil to lymphocyte ratio in patients with acute ST segment elevation myocardial infarction after percutaneous coronary intervention: a meta-analysis. *BMC Cardiovasc Disord*. 2018;18(1):75. doi:10.1186/s12872-018-0812-6

28. Al Aseri ZA, Habib SS, Marzouk A. Predictive value of high sensitivity C-reactive protein on progression to heart failure occurring after the first myocardial infarction. *Vasc Health Risk Manag*. 2019;15:221–227. doi:10.2147/VHRM.S198452

29. Xu M, Yan L, Xu J, Yang X, Jiang T. Predictors and prognosis for incident in-hospital heart failure in patients with preserved ejection fraction after first acute myocardial infarction: an observational study. *Medicine*. 2018;97(24):e11093. doi:10.1097/MD.0000000000011093

30. Stumpf C, Sheriff A, Zimmermann S, et al. C-reactive protein levels predict systolic heart failure and outcome in patients with first ST-elevation myocardial infarction treated with coronary angioplasty. *Arch Med Sci*. 2017;13(5):1086–1093. doi:10.5114/aoms.2017.69327

31. Varney JA, Dong VS, Tsao T, et al. COVID-19 and arrhythmia: an overview. *J Cardiol*. 2022;79(4):468–475. doi:10.1016/j.jjcc.2021.11.019

32. Kochi AN, Tagliari AP, Forleo GB, Fassini GM, Tondo C. Cardiac and arrhythmic complications in patients with COVID-19. *J Cardiovasc Electrophysiol*. 2020;31(5):1003–1008. doi:10.1111/jce.14479

33. Zhang Z, Yao L, Yang J, Wang Z, Du G. PI3K/Akt and HIF-1 signaling pathway in hypoxia-ischemia (Review). *Mol Med Rep*. 2018;18 (4):3547–3554. doi:10.3892/mmr.2018.9375

34. Liu X, Li M, Chen Z, et al. Mitochondrial calpain-1 activates NLRP3 inflammasome by cleaving ATP5A1 and inducing mitochondrial ROS in CVB3-induced myocarditis. *Basic Res Cardiol*. 2022;117(1):40. doi:10.1007/s00395-022-00948-1

35. Stewart RM, Fabian TC, Croce MA, Pritchard FE, Minard G, Kudsk KA. Is resection with primary anastomosis following destructive colon wounds always safe? *Am J Surg*. 1994;168(4):316–319. doi:10.1016/S0002-9610(05)80156-4

36. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell*. 2019;1(5):206–215. doi:10.1038/s42256-019-0048-x

37. Nohara Y, Iihara K, Nakashima N. Interpretable machine learning techniques for causal inference using balancing scores as meta-features. *Eng Med Biol Soc*. 2018;2018:4042–4045.

38. Figueroa Barraza J, López Droguett E, Martins MR. Towards interpretable deep learning: a feature selection framework for prognostics and health management using deep neural networks. *Sensors*. 2021;21(17). doi:10.3390/s21175888