THE ANATOLIAN JOURNAL OF CARDIOLOGY

Official journal of the TURKISH SOCIETY OF CARDIOLOGY

# A Machine Learning-Based Prognostic Risk Scoring Model for Patients with Coronary Artery Disease Admitted to the Intensive Care Unit Using Multi-Source Structured Clinical Data From the Information Mart for Intensive Care IV Database

## ABSTRACT

**Background:** Critical coronary artery disease (CAD) exhibits significant prognostic heterogeneity. Traditional scoring has limited efficacy in the early identification of high-risk patients. To develop a validated machine learning (ML)-based prognostic risk scoring model for critical CAD to predict adverse outcome risk based on multi-source structured clinical data from the medical information mart for intensive care IV (MIMIC-IV) database.

**Methods:** A total of 5472 critically ill patients with CAD meeting the operational definitions of this study were retrospectively enrolled from the MIMIC-IV database and randomly divided into a training set and a test set at a 7 : 3 ratio. Fifty-two structured clinical indicators collected within the time window from 6 hours before intensive care unit (ICU) admission to 24 hours after admission were extracted. Least absolute shrinkage and selection operator regression was applied in the training set to screen core variables based on 1-year all-cause mortality. Multiple ML algorithms were utilized to construct prognostic prediction models, with 1-year all-cause mortality as the primary outcome and 28-day in-hospital mortality as well as CAD-related ICU readmission rate as secondary outcomes. All model parameters and classification thresholds were predefined before performance evaluation in the test set. The test set was excluded from feature screening, parameter tuning, and threshold optimization procedures. Model performance was assessed via the area under the receiver operating characteristic curve (AUC of ROC), paired DeLong test, calibration curves, Brier score, calibration intercept, calibration slope, and decision curve analysis. The light gradient boosting machine (LightGBM) model with the highest AUC value was interpreted using SHapley additive exPlanations (SHAP). Additionally, sensitivity analysis with 5 sets of multiple imputations was performed.

**Results:** The 1-year all-cause mortality, 28-day in-hospital mortality, and CAD-related ICU readmission rate of the study population were 14.2%, 5.6%, and 9.1%, respectively. In the test set, the LightGBM model yielded AUC values of 0.896 and 0.898 for predicting 1-year all-cause mortality and 28-day in-hospital mortality, which were higher than those of other models. Nevertheless, paired DeLong tests revealed no statistically significant differences when compared with the logistic regression model ($P = .062$ and $P = .642$, respectively). The LightGBM model exhibited limited predictive performance for CAD-related ICU readmission rate (AUC = 0.690), with no significant difference in AUC relative to the logistic regression model ($P = .238$). SHapley additive exPlanations analysis identified the Charlson comorbidity index, history of coronary artery bypass grafting, and blood urea nitrogen as the top contributing variables of the LightGBM model. Thresholds derived from SHAP dependence plots were only interpreted as exploratory and model-dependent findings.

**Conclusion:** The LightGBM model established in this study exhibited favorable performance for predicting mortality risk among critically ill patients with CAD in internal validation and was interpretable via the SHAP approach. It can provide quantitative references for early risk stratification and individualized clinical management of critically ill CAD patients in the ICU. However, only internal validation was conducted in the present study, and the model demonstrated suboptimal predictive efficacy and calibration for CAD-related ICU readmission. Therefore, external, multicenter, and prospective validations are required prior to its clinical implementation.

**Keywords:** Coronary artery disease, machine learning, MIMIC-IV database, risk scoring model

Copyright@Author(s) - Available online at anatoljcardiol.com.
Content of this journal is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ORIGINAL INVESTIGATION

**Lingpei Liu**[1,2]
**Jinlei Zheng**[2]
**Xiuyang Li**[1]

[1]*Department of Big Data in Health Science, School of Public Health, Zhejiang University, Hangzhou, Zhejiang, China*
[2]*Department of General Practice, Affiliated Jinhua Hospital, Zhejiang University School of Medicine, Jinhua, Zhejiang, China*

**Corresponding author:**
Xiuyang Li
✉ lixiuyang@zju.edu.cn

Received date: February 9, 2026
Accepted date: August 28, 2026
**Available Online Date:**
November 13, 2026

**Cite this article as:** Liu L, Zheng J, Li X. A machine learning-based prognostic risk scoring model for patients with coronary artery disease admitted to the intensive care unit using multi-source structured clinical data from the Information Mart for Intensive Care IV database. *Anatol J Cardiol.* 2026;XX(X):X-X.

DOI: 10.14744/AnatolJCardiol.2026.6296

## INTRODUCTION

Coronary artery disease (CAD) remains a leading global cause of cardiovascular-related mortality and disability, pathologically characterized by atherosclerotic plaque formation in the coronary arteries, which can lead to myocardial ischemia, angina, myocardial infarction, and sudden cardiac death.[1,2] Indicated by global epidemiological data, CAD accounts for approximately 7 million deaths and 129 million disability-adjusted life years lost annually, highlighting its severity as a major public health concern.[2,3] Recent advancements, including revascularization techniques such as percutaneous coronary intervention (PCI) and coronary artery bypass grafting (CABG), and pharmacotherapies centered on antiplatelet agents and statins, have collectively driven optimizations of CAD management.[4,5] However, these treatments face challenges: PCI and CABG are associated with risks of restenosis, thrombosis, and bleeding, and other complications; long-term pharmacotherapy, though effective for disease control, is accompanied by increased bleeding risks or drug resistance.[6-8] Furthermore, critical CAD patients admitted to the intensive care unit (ICU) often present with complex pathophysiological states and multiple comorbidities, leading to highly heterogeneous treatment responses and clinical outcomes, with a severely high risk of adverse events.[9,10] Thus, the early identification of high-risk patients within the ICU is paramount for formulating targeted interventions, rationally allocating medical resources, and improving clinical outcomes.

In the ICU setting, scoring systems such as the acute physiology and chronic health evaluation (APACHE) and the sequential organ failure assessment (SOFA) are commonly used tools for assessing critically ill patients.[11,12] These systems provide clinicians with valuable references by comprehensively evaluating a patient's physiological status and organ function.[13] Nevertheless, these conventional scoring systems still possess certain limitations. For one thing, APACHE and SOFA are general assessment tools for critically ill patients rather than instruments tailored to predict long-term prognosis in patients with CAD. For another, such scoring systems rely on prefixed weights and a restricted panel of variables, making it difficult to comprehensively integrate multi-source structured clinical data, vital signs, and laboratory parameters extracted from electronic health records. When confronted with complicated pathophysiological conditions and multiple comorbidities, these scoring systems may fail to adequately capture individualized risk disparities among critically ill CAD patients, and their performance for long-term prognostic prediction remains unsatisfactory.[14] Therefore, there is a pressing need to develop more targeted risk assessment models.

The emergence of machine learning (ML) offers a novel methodological pathway to overcome the limitations of traditional prediction models. In contrast to some conventional statistical models that rely on linear or additive relationships under default settings, ML algorithms can learn from high-dimensional electronic health records and handle complex nonlinear associations and interaction effects among multi-source structured variables, including demographic characteristics, prior treatments, organ function scores, vital signs, and laboratory parameters.[15] The value of ML in predicting outcomes for critically ill patients has been preliminarily validated. In a prognostic study focusing on patients with acute kidney injury admitted to the ICU, researchers developed multiple ML models integrating multidimensional data such as medication dosages and clinical indicators, which improved predictive performance for patients' prognostic outcomes.[16] Furthermore, several studies have analyzed patients with ischemic stroke by constructing an interpretable ML framework, suggesting that the integration of multi-source structured clinical variables can improve discriminative ability for predicting in-hospital mortality risk.[17] Nevertheless, for critically ill patients with CAD—a population burdened with underlying cardiovascular disorders, acute critical illness, and multiple comorbidities—further research is still needed to establish stable and interpretable prognostic prediction models using early structured clinical data collected upon ICU admission.

Given the demonstrated potential of ML in risk assessment in the ICU settings, this study aims to leverage the medical information mart for intensive care IV (MIMIC-IV) database to construct an ML-based prognostic model for critical CAD patients. This study strictly follows the ML modeling standards. While ensuring the independence of data preprocessing and model construction, this study aims to develop and internally validate a risk stratification tool with promising predictive performance, so as to provide a quantitative reference for the early clinical identification of high-risk patients with CAD.

### HIGHLIGHTS

- In the test set, the light gradient boosting machine (LightGBM) model achieved favorable discriminative performance for predicting 1-year all-cause mortality and 28-day in-hospital mortality among critically ill patients with coronary artery disease, yielding the highest area under the curve (AUC) values across all candidate models. However, paired DeLong tests revealed no statistically significant AUC differences relative to the logistic regression model.
- SHapley additive exPlanations analyses demonstrated that the Charlson comorbidity index, prior coronary artery bypass grafting surgery history, and blood urea nitrogen were key predictors with substantial predictive contributions within the LightGBM model.

## METHODS

### Data Source and Study Population

Medical information mart for intensive care IV database (version 2.2) contains de-identified clinical data of critically ill patients admitted to the Medical Center between 2008 and 2019. This secondary analysis was performed using a publicly available de-identified database without access to identifiable individual information; therefore, additional ethical approval and informed consent were not required.[18]

First, patients with CAD-related International Classification of Diseases (ICD) codes were identified from the MIMIC-IV database. Coronary artery disease was defined according to the ICD codes: ICD-9 codes ranging from 410.xx to 414.xx, and ICD-10 codes I20.x-I22.x, I25.x. To increase the likelihood that CAD constituted the primary clinical issue of the index hospitalization, only patients with the above CAD codes listed among their top 3 discharge diagnoses were retained. This strategy was designed to mitigate potential misclassification caused by historical CAD or incidental CAD coding; however, it may reduce the generalizability of the findings to all patients with CAD.

In this study, critically ill CAD was operationally defined as patients aged ≥ 18 years who were admitted to the ICU for the first time during the hospitalization, had an ICU stay of at least 24 hours, and carried the aforementioned CAD-related ICD codes within their top 3 discharge diagnoses. This definition was solely applied for participant screening and did not serve as an independent clinical severity grading criterion.

Initial screening yielded 42264 patients based on the criteria of age, first ICU admission, and ICU length of stay. Among them, 6777 patients met the operational definition of critically ill CAD after screening for CAD-related ICD codes in the top 3 discharge diagnoses. An additional 1305 patients without accessible electrocardiogram (ECG) records from 6 hours before to 24 hours after ICU admission were further excluded, leaving a final cohort of 5472 patients for analysis (Figure 1).

## Predictor Collection and Selection

Using the MIMIC-IV database, clinical data were extracted from 6 hours before to 24 hours after ICU admission, covering 52 variables including demographics, prior treatments, organ-function scores, vital signs, laboratory values, and electrocardiographic parameters (Supplementary Table 1). For continuous variables, the maximum or minimum value of each variable within the predefined time window was extracted according to clinical relevance. For variables with multiple serial measurements, the corresponding extremum was defined as the single most abnormal measurement recorded in this window, which was used to reflect aberrant physiological status in the early phase of ICU admission.[19]

All variable selection procedures were performed strictly only in the training set to prevent information leakage from the testing set. At the variable-screening stage, a logistic-regression model was constructed with the least absolute shrinkage and selection operator (LASSO) regression. One-year all-cause mortality served as the outcome, and the collected clinical indices were entered as predictors. By adding an $L_1$-penalty to the loss function, LASSO simultaneously performs feature selection and shrinks redundant coefficients to mitigate multicollinearity. Model fitting was performed using the "glmnet" R package, and the penalty parameter $\lambda$ was determined via 10-fold cross-validation. To obtain a

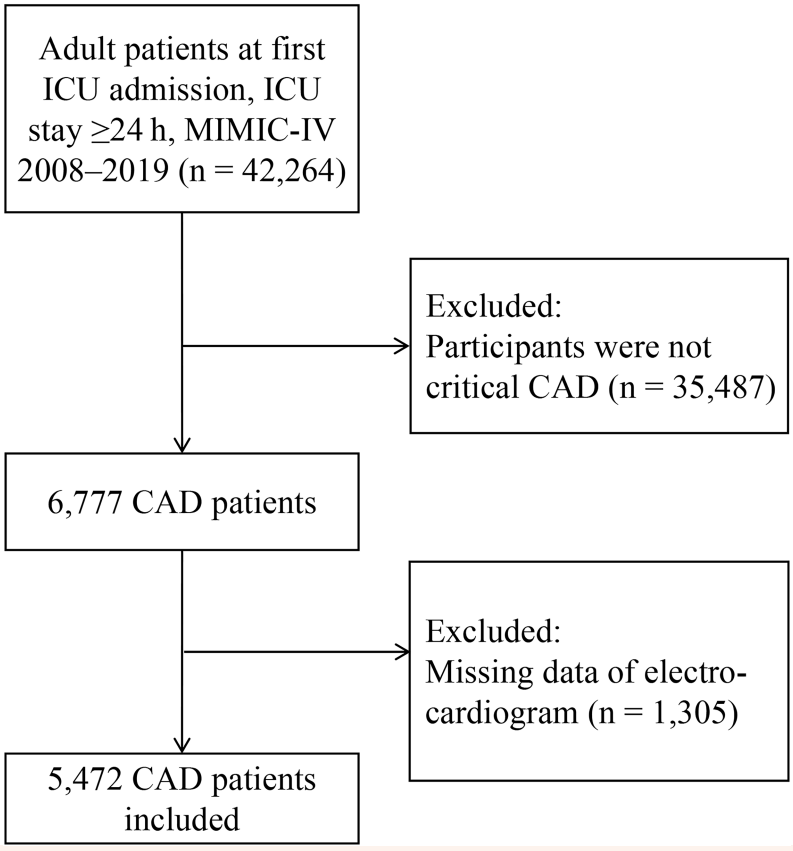



**Figure 1. Study flowchart of participant recruitment.**

more parsimonious and stable set of variables, the $\lambda$ value corresponding to the 1-standard-error rule ($\lambda_{1se}$) was adopted for feature selection in this study. Variables retaining non-zero coefficients were retained as candidate features, and their absolute coefficient values were plotted to generate an importance profile that visually illustrates the coefficient direction and relative weight of selected variables in the LASSO model. One-year all-cause mortality was adopted as the endpoint for feature selection in this study, primarily because this outcome yielded a relatively large number of events and could comprehensively reflect the overall disease burden among critically ill patients with CAD. The identical panel of 13 predictors was retained for models of 28-day in-hospital mortality and CAD-related ICU readmission. This design aimed to evaluate the predictive performance of the same parsimonious variable set across distinct endpoints and mitigate the risk of overfitting induced by repeated variable screening for each secondary outcome. It should be noted that this variable panel was selected based on the endpoint of 1-year all-cause mortality. Therefore, its application to the 2 secondary outcomes merely serves to assess the predictive capacity of a unified variable framework, rather than representing the optimal feature combination for predicting 28-day in-hospital mortality or CAD-related ICU readmission.

### Outcome Variables

The primary outcome of this study was all-cause mortality within 1 year after ICU admission. Secondary outcomes consisted of 28-day in-hospital mortality after ICU admission and CAD-related ICU readmission. Twenty-eight-day in-hospital mortality was defined as death occurring within 28 days of ICU admission during the hospitalization. Coronary artery disease–related ICU readmission was defined as the first subsequent hospitalization with ICU admission recorded in MIMIC-IV version 2.2 after discharge from the hospitalization, where the top 3 discharge diagnoses of this subsequent hospital stay contained the predefined CAD ICD codes. Follow-up commenced upon discharge from the hospitalization, and all available records of subsequent hospitalizations in MIMIC-IV 2.2 were retrieved. Patients without a recorded CAD-related ICU readmission by the end of available database records were classified as having no observed readmission event within the database. Since MIMIC-IV cannot capture hospitalizations occurring at external medical institutions, this outcome only reflects CAD-related ICU readmissions observable within the single-center database and does not equate to the true readmission rate of the study cohort.

### Statistical Analysis and Model Development

The dataset was randomly split into a training set and a test set at a 7 : 3 ratio. To strictly prevent data leakage, all preprocessing and modeling procedures—including missing value imputation, LASSO feature selection, model fitting, hyperparameter tuning, and classification threshold determination—were conducted exclusively on the training set. The test set was only reserved for the final performance evaluation. All model hyperparameters were fixed in advance prior to test set assessment; no hyperparameter tuning or classification threshold optimization was performed using data from the test set.

After the 13 predictive variables were finalized, 9 models were separately constructed to predict prognostic risks in critically ill patients with CAD, including logistic regression, Random Forest, support vector machine, extreme gradient boosting, light gradient boosting machine (LightGBM), adaptive boosting, Naive Bayes, *K*-Nearest Neighbors, and categorical boosting (CatBoost).[20] Classification thresholds were predefined based on the event rates of each endpoint in the training set, with cutoffs of 0.18, 0.06, and 0.08 adopted for 1-year all-cause mortality, 28-day in-hospital mortality, and CAD-related ICU readmission, respectively, and applied uniformly across all models. The aforementioned thresholds were only used to calculate accuracy, sensitivity, and specificity.

The primary metric for evaluating model discriminative performance was the area under the receiver operating characteristic curve (AUC of ROC), alongside supplementary reporting of accuracy, sensitivity, and specificity. The 95% CIs for the AUC of each model were estimated via 2000 bootstrap resamples. Pairwise comparisons of AUC between the LightGBM and logistic regression models in the test set were performed using the paired DeLong test. Model calibration was assessed using calibration plots, the Brier score, calibration intercept, and calibration slope. Decision curve analysis (DCA) was utilized to quantify the net clinical benefit across a range of risk threshold probabilities. To enhance the interpretability of model outputs, SHapley additive exPlanations (SHAP) was applied to interpret the LightGBM model. Variable cutoffs derived from SHAP dependence plots serve only as exploratory signals specific to the model and do not constitute independently validated clinical thresholds.

Data were extracted from MIMIC-IV 2.2 with structured query language, processed and analyzed in R 4.4.1. During data preprocessing, single imputation using predictive mean matching was performed to construct the primary analytic dataset for the main analysis. To quantify uncertainty arising from missing data handling, a multiple imputation sensitivity analysis based on 5 imputed datasets was further conducted. The original split of training and test sets was retained unchanged, and multiple imputation by chained equations was adopted to generate 5 complete datasets with 10 iterations per imputation cycle. Imputation models were estimated exclusively on training set data without incorporating test set records to derive imputation parameters; missing values in the test set were imputed using the imputation framework trained on the training cohort, which prevents leakage of test-set information into preprocessing and model training. Subsequently, models were separately fitted on the training subset of each imputed dataset, and predicted probabilities were generated on the corresponding test subsets. The predicted probabilities of each individual across the 5 imputed test datasets were averaged, and the averaged probabilities were used to evaluate model performance on the test set.

Continuous variables were first examined for normality with Q–Q plots and Shapiro–Wilk tests. Normally distributed data are presented as mean ± SD and compared with 1-way ANOVA (≥3 groups) or *t*-test (2 groups). Non-normally distributed variables are summarized as median (interquartile range) and analyzed with the Kruskal–Wallis test (≥3 groups) or Mann–Whitney *U*-test (2 groups). Categorical variables are reported as counts and percentages, and intergroup differences were tested with the $\chi^2$ test. All tests were 2-sided, and $P < .05$ was considered statistically significant.

In addition, to explore selection bias potentially introduced by missing ECG records, baseline clinical characteristics as well as the incidence of 1-year all-cause mortality, 28-day in-hospital mortality, and CAD-related ICU readmission were further compared between patients with available ECG data and those without ECG recordings. The aforementioned statistical methods were applied for intergroup comparisons.

## RESULTS

### Least Absolute Shrinkage and Selection Operator Selection of Predictors and Baseline Analysis

During the variable selection phase, a total of 13 predictors with non-zero regression coefficients were screened out via LASSO regression analysis from 52 clinical indicators extracted from the MIMIC-IV database, which served as candidate predictors for subsequent model construction. These predictors were: CABG, Charlson comorbidity index (CCI), maximum anion gap (AG), SOFA score, maximum prothrombin time (PT), neutrophil-to-lymphocyte ratio (NLR), maximum respiratory rate, minimum pulse oxygen saturation, maximum blood urea nitrogen (BUN), weight, maximum partial thromboplastin time (PTT), QRS duration, and maximum heart rate (HR) (Figure 2).

A total of 5472 CAD patients were enrolled. The 1-year mortality was 14.2%, the 28-day in-hospital mortality was 5.6%, and the CAD-related ICU readmission rate was 9.1% (Supplementary Table 2). Patients were randomly split 7 : 3 into a training set and a testing set; baseline characteristics are listed in Supplementary Table 2. Apart from AG ($P < .05$), no significant between-set differences were found for any other variables ($P > .05$), indicating satisfactory overall comparability between the 2 groups.

In addition, supplementary comparisons between patients with available ECG records and those without ECG data revealed that patients with missing ECG recordings had a higher CCI (5.55 vs. 4.91), along with elevated rates of 1-year all-cause mortality and 28-day in-hospital mortality (18.9% vs. 14.2%; 8.0% vs. 5.6%). In contrast, the CAD-related ICU readmission rates were comparable between the 2 groups (9.9% vs. 9.1%) (Supplementary Table 3).

### Comparison of Prediction Performance of Different Machine Learning Models for 1-Year Mortality Risk of Critical Coronary Artery Disease Patients

Table 1 summarizes the predictive performance of each ML model for 1-year all-cause mortality risk among critically ill patients with CAD in both the training and test sets. The predefined hyperparameters of all models and classification thresholds corresponding to the 3 endpoints are presented in Supplementary Table 4. The LightGBM model yielded an AUC of 0.921 in the training set and 0.896 in the test set; its test-set AUC was numerically higher than that of logistic

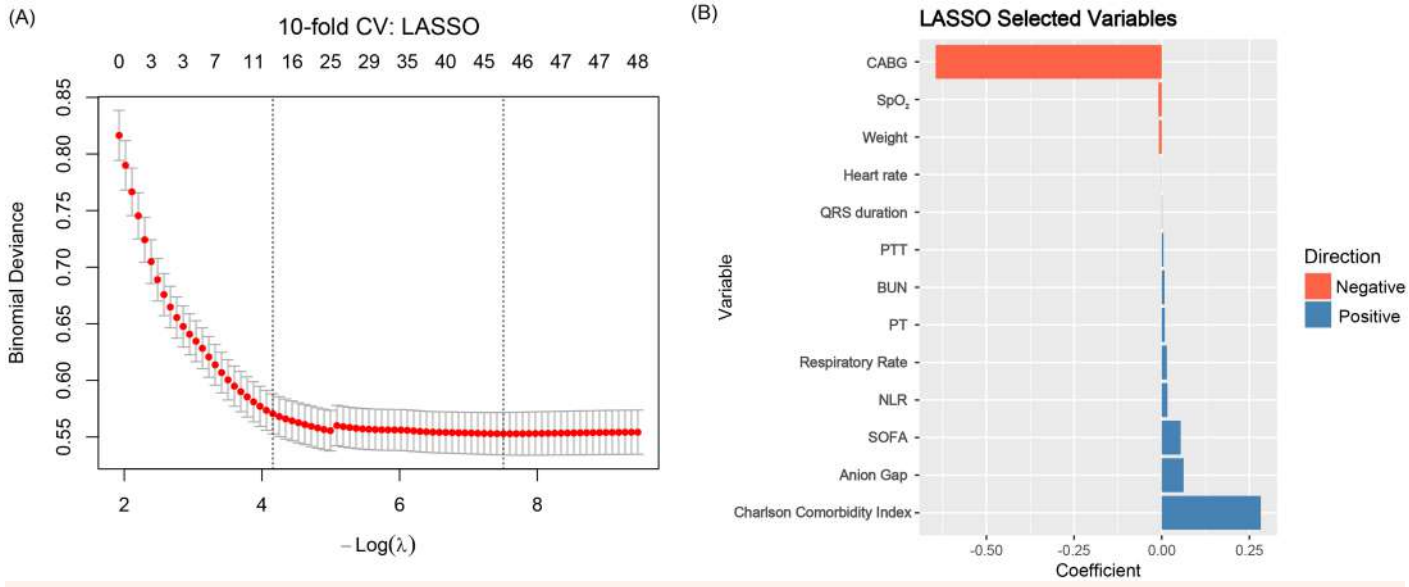



**Figure 2. LASSO feature selection and coefficient visualization. (A) Ten-fold cross-validation curve used to determine the penalty parameter $\lambda$. The vertical dotted lines indicate $\lambda_{min}$ and $\lambda_{1se}$. (B) Bar plot of the coefficients of the 13 variables selected at $\lambda_{1se}$, with colors indicating coefficient directions. AG, anion gap; BUN, blood urea nitrogen; CABG, coronary artery bypass grafting; CCI, Charlson comorbidity index; HR, heart rate; NLR, neutrophil-to-lymphocyte ratio; PT, prothrombin time; PTT, partial thromboplastin time; QRS, QRS duration; RR, respiratory rate; SOFA, Sequential Organ Failure Assessment; $SpO_2$, pulse oxygen saturation; WT, weight.**

**Table 1. Performance Comparison of ML Models for Predicting 1-Year Mortality Risk in CAD Patients in the Training and Testing Sets**

| Model | AUC | Accuracy | Sensitivity | Specificity |
|---|---|---|---|---|
| Train | | | | |
| AdaBoost | 0.904 (0.892, 0.915) | 0.858 | 0.708 | 0.883 |
| Naive Bayes | 0.865 (0.850, 0.881) | 0.850 | 0.642 | 0.884 |
| CatBoost | 0.881 (0.867, 0.894) | 0.773 | 0.836 | 0.762 |
| *K*-nearest neighbors | 0.895 (0.883, 0.908) | 0.847 | 0.757 | 0.862 |
| LightGBM | 0.921 (0.910, 0.932) | 0.849 | 0.833 | 0.851 |
| Logistic regression | 0.884 (0.870, 0.899) | 0.836 | 0.763 | 0.848 |
| Random forest | 0.889 (0.873, 0.904) | 0.891 | 0.520 | 0.953 |
| Support vector machine | 0.904 (0.887, 0.921) | 0.919 | 0.711 | 0.954 |
| XGBoost | 0.899 (0.886, 0.912) | 0.769 | 0.866 | 0.753 |
| Test | | | | |
| AdaBoost | 0.862 (0.838, 0.886) | 0.840 | 0.665 | 0.869 |
| Naive Bayes | 0.866 (0.842, 0.891) | 0.851 | 0.691 | 0.877 |
| CatBoost | 0.877 (0.854, 0.900) | 0.766 | 0.854 | 0.752 |
| *K*-nearest neighbors | 0.879 (0.854, 0.904) | 0.844 | 0.764 | 0.858 |
| LightGBM | 0.896 (0.875, 0.917) | 0.832 | 0.820 | 0.834 |
| Logistic regression | 0.887 (0.864, 0.911) | 0.831 | 0.798 | 0.837 |
| Random forest | 0.875 (0.850, 0.900) | 0.882 | 0.528 | 0.940 |
| Support vector machine | 0.849 (0.816, 0.881) | 0.872 | 0.639 | 0.911 |
| XGBoost | 0.883 (0.860, 0.906) | 0.747 | 0.884 | 0.725 |

AUC, area under the receiver operating characteristic curve (95% CI in parentheses). Accuracy: overall prediction accuracy; sensitivity: true-positive rate for mortality events; specificity: true-negative rate for survival. The training set was used for feature selection and model development; the held-out test set was used for independent internal performance evaluation.
AUC, area under the curve; CAD, coronary artery disease; ML, machine learning; LightGBM, light gradient boosting machine; XGBoost, extreme gradient boosting.

regression (AUC = 0.887) and all other ML algorithms (Table 1, Figure 3 A). The paired DeLong test revealed an AUC difference of 0.0091 between LightGBM and logistic regression in the test set (95% CI: −0.0004 to 0.0186, $P = .062$), indicating no statistically significant discrepancy.

For the endpoint of 1-year all-cause mortality, the LightGBM model achieved a Brier score of 0.081 in the test set, with a calibration intercept of 0.170 (95% CI: −0.057 to 0.397) and a calibration slope of 1.179 (95% CI: 1.040-1.318) (Supplementary Table 5). The 95% CI of the calibration intercept crossed zero, while the calibration slope exceeded 1 and its 95% CI excluded 1. These findings suggest that model-predicted probabilities tended to cluster toward the mid-range, with a limited magnitude of discrimination across risk strata. Calibration curves and DCA are displayed in Supplementary Figures 1-2, demonstrating a moderate net clinical benefit of the model within certain ranges of threshold probability.

### Comparison of Prediction Performance for 28-Day in-Hospital Mortality and Coronary Artery Disease–Related Intensive Care Unit Readmission Risk by Different Machine Learning Models

This study examined the performance of the different ML models for predicting the 2 secondary outcomes in critical CAD patients. As shown in Table 2 and Figures 3 B-C, in the test set, the LightGBM model yielded an AUC of 0.898 (95% CI: 0.870-0.925) for predicting 28-day in-hospital mortality risk, while the AUC of the logistic regression model was 0.894 (95% CI: 0.862-0.926). The paired DeLong test showed an AUC difference of 0.0036 between the 2 models (95% CI: −0.0117 to 0.0190, $P = .642$) (Supplementary Table 6), indicating that although the AUC value of LightGBM was slightly higher, the difference versus logistic regression lacked statistical significance. All models exhibited limited overall discriminative performance when predicting CAD-related ICU readmission risk. The test-set AUC of the LightGBM model was 0.690 (95% CI: 0.651-0.728), compared with 0.673 (95% CI: 0.633-0.713) for logistic regression. The AUC difference was 0.0166 (95% CI: −0.0110 to 0.0442, $P = .238$) (Supplementary Table 6), with no statistically significant gap observed.

Quantitative calibration metrics of the LightGBM model for the 2 secondary endpoints are summarized in Supplementary Table 5. For 28-day in-hospital mortality, the Brier score, calibration intercept, and calibration slope were 0.043, 0.098 (95% CI: −0.296 to 0.493), and 1.032 (95% CI: 0.865-1.198), respectively. For CAD-related ICU readmission, the corresponding values were 0.094, −0.203 (95% CI: −0.675 to 0.269), and 0.777 (95% CI: 0.582-0.973). Corresponding calibration curves and DCA are presented in Supplementary Figures 3-6.

Sensitivity analysis based on 5 multiply imputed datasets demonstrated that the LightGBM model achieved AUC values of 0.897, 0.898, and 0.684 for 1-year all-cause mortality, 28-day in-hospital mortality, and CAD-related ICU

readmission, respectively, which were largely consistent with the primary analysis results. The corresponding AUCs of the logistic regression model were 0.887, 0.896, and 0.665 (Supplementary Table 7). These findings indicated that different missing data handling strategies did not substantially alter the ranking of models by AUC or the interpretation of core conclusions.

### SHapley Additive ExPlanations Analysis Results

To enhance the interpretability of the LightGBM prediction model, SHAP analysis was employed to quantify the contribution of each feature to the prediction of 1-year mortality risk in critical CAD patients. Figure 4 illustrates the global impact of each feature. The SHAP summary plot illustrated that the CCI yielded the largest mean absolute SHAP value, making it the predictor with the greatest contribution to the LightGBM model for predicting 1-year all-cause mortality risk (Figure 4 A). Further analyses revealed that a higher CCI, SOFA score, NLR, QRS duration, AG, and absence of prior CABG surgery were associated with elevated SHAP values, indicating increased predicted mortality risk generated by the model (Figure 4 B). Additionally, SHAP dependence analysis was conducted for all feature indicators to assess the contribution of each feature to the model output and its relationship with the actual feature values. Higher predicted 1-year all-cause mortality risk generated by the model was associated with the following conditions: no history of CABG surgery, CCI $\geq$ 5, AG $\geq$ 18 mEq/L, respiratory rate $\geq$ 27 breaths per minute, $SpO_2 \leq$ 86% or >96%, body weight $\geq$ 86 kg, QRS duration $\geq$ 122 ms, and HR $\geq$ 110 bpm (Supplementary Figures 7-19). The above cutoffs merely depict local nonlinear associations specific to the current dataset and the LightGBM algorithm and should not be regarded as clinically validated thresholds. For observations inconsistent with conventional clinical expectations—including elevated body weight, $SpO_2$ > 96%, and absence of prior CABG—cautious interpretation is required, with potential explanations encompassing residual confounding, treatment selection bias, heterogeneous measurement contexts, and survivor bias.

Furthermore, SHAP analysis was similarly applied to quantify the contribution of features to the prediction of the 2 secondary outcomes. As shown in Supplementary Figures 20-21, for the prediction of 28-day in-hospital mortality risk, elevated SHAP values were observed in patients with a higher AG, SOFA score, CCI, NLR, and no prior history of CABG surgery; these variables represented the top contributors to predicting this endpoint (Supplementary Figure 20). Regarding CAD-related ICU readmission risk, higher SHAP values were observed in patients with an elevated CCI, BUN, AG, PTT, and lower HR; these variables made substantial contributions to the LightGBM predictive model (Supplementary Figure 21).

## DISCUSSION

This study, based on the multi-source structured clinical data from the MIMIC-IV database, developed and evaluated multiple ML models for predicting the prognosis of critical CAD patients. The results demonstrated that the LightGBM model achieved numerically higher test-set AUC values of 0.896 and 0.898 for 1-year all-cause mortality and 28-day in-hospital mortality, respectively. Nevertheless, pairwise DeLong comparisons against logistic regression failed to reach statistical significance for both endpoints. The model exhibited relatively limited discriminative capacity for CAD-related ICU readmission risk (AUC = 0.690). SHAP analyses identified the CCI, history of CABG surgery, NLR, and AG as the clinical features with the largest predictive contributions. These findings indicate that the LightGBM model built on structured early ICU admission clinical variables can serve as a potential quantitative reference for risk stratification in critically ill patients with CAD, yet its clinical utility remains to be further validated via external cohorts.

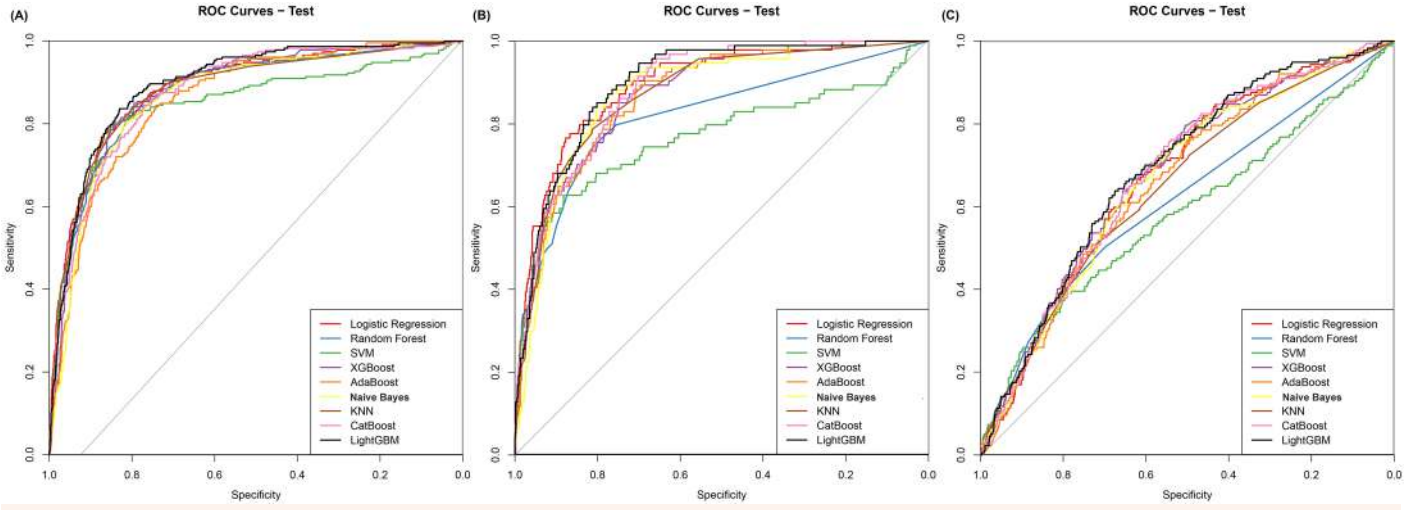



**Figure 3. ROC curves for predicting clinical outcomes. ROC curves for the predictive models for 3 outcomes: (A) primary outcome: 1-year mortality; (B) secondary outcome 1: 28-day in-hospital mortality; (C) secondary outcome 2: CAD-related ICU readmission. AUC indicates the discriminatory ability of each model. AUC, area under the curve; CAD, coronary artery disease; ICU, intensive care unit; ROC, receiver operating characteristic.**

**Table 2. Predictive Performance of ML Models for 28-Day in-Hospital Mortality and CAD-Related ICU Readmission**

| Model | 28-Day in-Hospital Mortality | | | | CAD-Related ICU Readmission | | | |
|---|---|---|---|---|---|---|---|---|
| | AUC | Accuracy | Sensitivity | Specificity | AUC | Accuracy | Sensitivity | Specificity |
| AdaBoost | 0.866 (0.833, 0.899) | 0.833 | 0.691 | 0.842 | 0.682 (0.643, 0.720) | 0.767 | 0.395 | 0.812 |
| Naive Bayes | 0.874 (0.841, 0.907) | 0.840 | 0.745 | 0.846 | 0.666 (0.626, 0.707) | 0.750 | 0.384 | 0.794 |
| CatBoost | 0.885 (0.857, 0.914) | 0.865 | 0.649 | 0.878 | 0.677 (0.638, 0.716) | 0.738 | 0.424 | 0.776 |
| *K*-nearest neighbors | 0.872 (0.837, 0.907) | 0.812 | 0.777 | 0.814 | 0.650 (0.608, 0.692) | 0.699 | 0.508 | 0.722 |
| LightGBM | 0.898 (0.870, 0.925) | 0.817 | 0.830 | 0.817 | 0.690 (0.651, 0.728) | 0.661 | 0.644 | 0.663 |
| Logistic regression | 0.894 (0.862, 0.926) | 0.792 | 0.819 | 0.791 | 0.673 (0.633, 0.713) | 0.664 | 0.599 | 0.672 |
| Random forest | 0.800 (0.749, 0.852) | 0.937 | 0.372 | 0.971 | 0.601 (0.560, 0.643) | 0.890 | 0.017 | 0.995 |
| Support vector machine | 0.768 (0.705, 0.832) | 0.882 | 0.585 | 0.900 | 0.583 (0.534, 0.632) | 0.667 | 0.452 | 0.694 |
| XGBoost | 0.868 (0.833, 0.904) | 0.865 | 0.670 | 0.877 | 0.679 (0.639, 0.719) | 0.685 | 0.565 | 0.700 |

Values are presented as performance metrics derived from the test dataset. The 95% CI for AUC values were obtained using 2000 bootstrap replicates.
AUC, area under the curve; CAD, coronary artery disease; ICU, intensive care unit; LightGBM, light gradient boosting machine; ML, machine learning; XGBoost, extreme gradient boosting.

The LightGBM model exhibited strong discriminative performance for 1-year all-cause mortality among critically ill patients with CAD, yet the AUC difference versus logistic regression was merely 0.0091, and the paired DeLong test did not reach statistical significance ($P = .062$). Accordingly, current evidence fails to demonstrate that LightGBM is statistically superior to logistic regression. Gradient boosting tree algorithms are capable of capturing intricate nonlinear associations and variable interactions, which may facilitate modeling complex interplays among systemic inflammatory response, multiple organ dysfunction, comorbidity burden, and other clinical factors in critically ill CAD patients.[21-23] Notably, for prediction models constructed from structured clinical data, model performance should not be judged solely by AUC magnitude. Instead, a comprehensive evaluation incorporating pairwise statistical comparisons, calibration

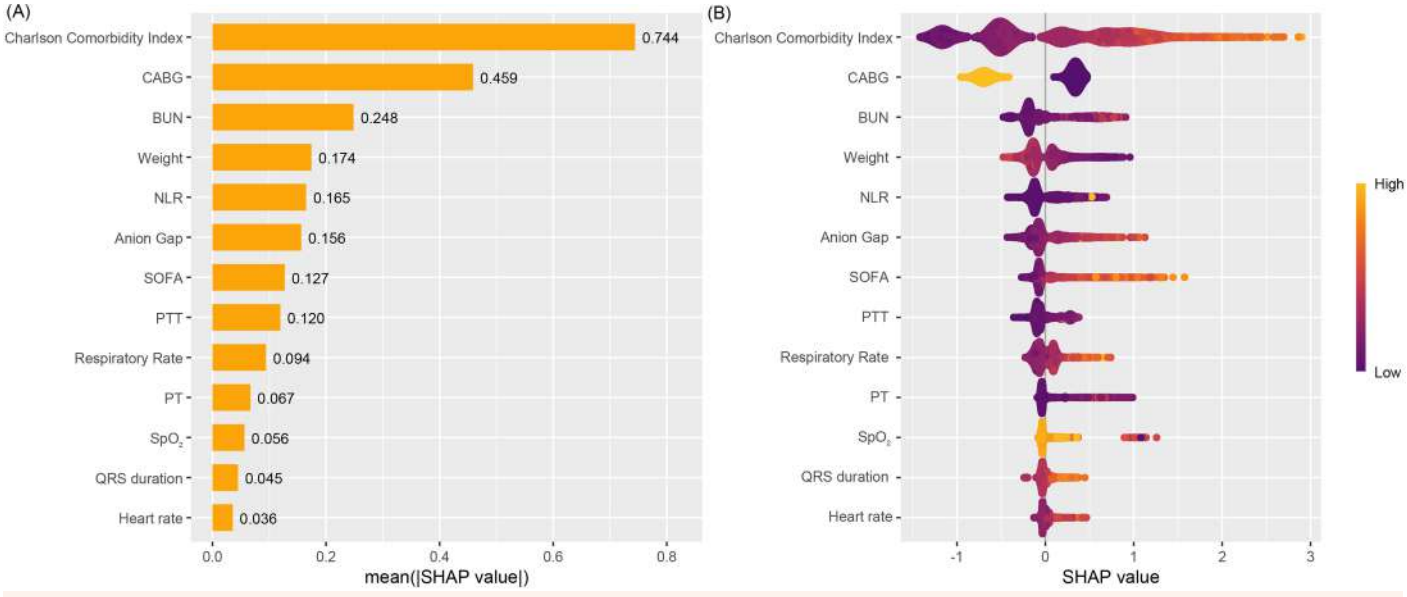



**Figure 4. SHAP feature importance and summary for 1-year mortality prediction. (A) SHAP feature importance bar plot showing the contribution of each feature to the prediction of one-year mortality. Larger values indicate a greater contribution to the model prediction. (B) SHAP bee swarm plot, where each point represents an individual sample. The X-axis shows the SHAP value for the feature, and the color represents the actual feature value (yellow for high, purple for low), illustrating how feature values relate to mortality risk and their distribution. AG, anion gap; BUN, blood urea nitrogen; CABG, coronary artery bypass grafting; CCI, Charlson comorbidity index; HR, heart rate; NLR, neutrophil-to-lymphocyte ratio; PT, prothrombin time; PTT, partial thromboplastin time; QRS, QRS duration; RR, respiratory rate; SHAP, SHapley additive exPlanations; SOFA, sequential organ failure assessment; $SpO_2$, pulse oxygen saturation; WT, weight.**

metrics, decision curve-derived net clinical benefit, and external validation results is recommended.[24-26]

In terms of calibration, the LightGBM model had a calibration intercept of 0.170 for 1-year all-cause mortality, with its 95% CI crossing zero; the calibration slope was 1.179 and its 95% CI excluded 1. This pattern indicated a compressed distribution of model-predicted probabilities, meaning interindividual differences in prognostic risk might be underestimated. Decision curve analysis revealed modest net clinical benefit within certain ranges of risk threshold probability. Nevertheless, only internal validation was performed in the present study, so future assessments of calibration performance and recalibration, if required, are warranted using independent external cohorts. Furthermore, an important limitation regarding model comparisons merits clarification: this study did not conduct direct head-to-head comparisons between the LightGBM model and predictions derived from standalone APACHE or SOFA scores, as the SOFA score itself was incorporated as 1 predictor within the LightGBM framework. Accordingly, the current findings cannot substantiate that LightGBM outperforms conventional critical illness scoring systems such as APACHE and SOFA. Direct head-to-head analyses within identical study populations are needed in subsequent research.

The predictive performance of the LightGBM model differed notably between the 2 secondary outcomes. Since 28-day in-hospital mortality and 1-year all-cause mortality partially share overlapping pathophysiological underpinnings, including multiple organ dysfunction and systemic inflammatory response, the predictor set selected based on long-term all-cause mortality maintained favorable discriminative and calibrated performance for short-term death prediction.[22,23] Nonetheless, the lack of statistically significant differences between LightGBM and logistic regression indicated that its marginal numerical superiority failed to demonstrate a definitive improvement in predictive efficacy. By contrast, the LightGBM model exhibited relatively weak discrimination and calibration for CAD-related ICU readmission. The calibration slope of the readmission model was below 1, which implied overly extreme predicted probabilities and hinted at mild overfitting risk. Uniform adoption of the identical 13 predictors preserved consistency across models for distinct endpoints and mitigated overfitting induced by repeated variable selection for secondary outcomes with smaller event counts. However, these predictors were screened targeting 1-year all-cause mortality without tailored optimization for readmission events, so they could not fully capture core determinants of hospital readmission. Additionally, readmission risk is jointly shaped by in-hospital clinical status, post-discharge treatment, and social contextual factors. Accordingly, results for the CAD-related ICU readmission model should be interpreted as exploratory signals only.

Coronary artery disease–related ICU readmission arises from a combination of in-hospital illness trajectories and post-discharge management. Beyond cardiac triggers such as progressive heart failure and recurrent myocardial ischemia, noncardiac acute illnesses, medication adherence after discharge, chronic disease self-management, and outpatient follow-up attendance all modify readmission risk in CAD patients.[27-29] However, the MIMIC-IV database primarily archives inpatient clinical data and lacks granular post-discharge follow-up records, outpatient adherence metrics, and socioeconomic covariates. This data limitation hinders the comprehensive capture of the multifactorial drivers of readmission and largely accounts for the model's modest predictive power. Accordingly, the predictive findings for CAD-related ICU readmission remain exploratory and require further external validation using independent cohorts supplemented with comprehensive post-hospitalization information.

SHapley additive exPlanations analysis identified the clinical variables with the greatest predictive contributions to the LightGBM model for prognostic risk estimation in critically ill patients with CAD. Among these, the CCI represented the predictor with the largest contribution to the model when forecasting 1-year all-cause mortality. As a composite measure quantifying the cumulative burden of chronic diseases (e.g., diabetes, chronic kidney disease, peripheral vascular disease, and chronic obstructive pulmonary disease), a high CCI signifies reduced physiological reserve and increased systemic vulnerability.[30] In the context of CAD, these comorbidities exacerbate the risk of adverse outcomes through various pathways. For instance, diabetes accelerates atherosclerotic progression and impairs coronary microcirculation by promoting endothelial dysfunction, oxidative stress, and chronic inflammation.[31] Chronic kidney disease not only acts synergistically with traditional cardiovascular risk factors, but its unique pathological states, such as uremic toxin accumulation, calcium–phosphorus metabolism disorders, and anemia, further aggravate myocardial fibrosis and vascular calcification, collectively increasing the risk of cardiac death and heart failure.[32-34] Accordingly, the substantial contribution of the CCI within the model suggests that systemic health status reflected by overall comorbidity burden constitutes a key component of prognostic risk in critically ill patients with CAD.

A prior history of CABG exhibited a protective predictive signal in the model, which aligns broadly with previous evidence that CABG may improve long-term outcomes in selected patients with multivariable CAD. Prior studies have demonstrated that CABG is associated with lower long-term mortality risk relative to PCI among certain patients with multivariable coronary lesions.[35,36] Nevertheless, the present analysis was not designed as a comparative effectiveness study of different revascularization strategies. A history of prior CABG may simultaneously reflect treatment selection bias, anatomical severity of coronary disease, healthcare access, and survivor effects. As such, the protective direction of this variable within the model cannot be interpreted as a direct causal protective effect of CABG in the current cohort. In addition, the AG and BUN also carried moderate predictive weight in the model. An elevated AG indicates excess unmeasured anions and acid-base disturbances, frequently occurring in the setting of renal dysfunction or

inadequate tissue perfusion.[37] Increased BUN levels are linked to multiple biological processes, including impaired renal function, altered volume status, accelerated protein catabolism, and neuroendocrine activation.[38] The predictive contributions of these biomarkers likely reflect the underlying roles of renal dysfunction, systemic perfusion status, and cardiorenal crosstalk in shaping prognostic risk for critically ill patients with CAD.[39]

Several limitations of this study should be noted. First, the study relied on a single-center retrospective database and only completed internal validation, leaving the generalizability of the model untested via multicenter prospective cohorts. As such, the current model should be regarded merely as an investigational risk-stratification tool rather than a direct aid for routine clinical decision-making. Second, the requirement for valid ECG recordings collected between 6 hours before ICU admission and 24 hours post-admission may have introduced selection bias. Supplementary comparative analyses revealed that patients without available ECG data carried a higher CCI, alongside elevated 1-year all-cause mortality and 28-day in-hospital mortality. This suggests the final analytical cohort had an overall lower baseline risk burden relative to patients excluded due to missing ECG records. Third, the predictive framework was built exclusively on static structured clinical variables captured at early ICU admission. Dynamic temporal fluctuations of physiological parameters during hospitalization, as well as unstructured data including imaging scans, continuous waveform signals, and free-text clinical notes, were not incorporated. This data restriction prevents full characterization of evolving disease trajectories and multifaceted clinical phenotypes. Fourth, all candidate predictors were selected using LASSO regression targeted at the primary outcome of 1-year all-cause mortality; the identical fixed variable panel was subsequently reused for the 2 secondary outcomes without endpoint-specific feature screening. Therefore, this set of predictors may not represent the optimal feature combination for forecasting 28-day in-hospital mortality or CAD-related ICU readmission independently. Fifth, CAD-associated ICU readmission was defined solely based on subsequent ICU hospitalizations documented within the MIMIC-IV single institution, failing to capture readmission events occurring at external healthcare facilities. Furthermore, outcome identification depended on discharge diagnosis coding, which carries inherent risks of outcome misclassification. Accordingly, the observed 9.1% readmission proportion reflects only CAD-related ICU revisit rates recorded in the database, rather than the true population-level readmission incidence. Finally, the MIMIC-IV database lacks granular coronary anatomical details, post-discharge medication adherence metrics, social support indicators, and long-term outpatient follow-up records. Absence of these out-of-hospital covariates substantially attenuates the model's capacity to predict complex readmission endpoints. Future research should conduct rigorous external validation on independent cohorts, perform recalibration where indicated, and integrate multicenter datasets, time-varying physiological signals, and comprehensive post-discharge clinical variables when feasible.

## CONCLUSION

This study constructed an ML prognostic risk model for critically ill patients with CAD based on multi-source structured clinical data collected in the early phase of ICU admission and completed its internal validation. The results demonstrated that the LightGBM model exhibited favorable discriminative performance for the outcomes of 1-year all-cause mortality and 28-day in-hospital mortality, and its predictions could be interpreted via SHAP analysis. Nevertheless, no statistically significant differences in AUC were detected between LightGBM and logistic regression, and the model showed limited predictive efficacy for CAD-related ICU readmission. Accordingly, the LightGBM model may serve as a potential investigational reference for early risk stratification of critically ill CAD patients, yet it is not suitable for direct clinical decision-making at present. Since only internal validation was performed in the current analysis, the generalizability of the model across diverse medical centers and patient populations requires further verification using multicenter prospective external cohorts, with recalibration conducted if necessary.

**Ethics Committee Approval:** The dataset is a derivative dataset of MIMIC-IV and thus no new patient data was collected. The ethics approval of the dataset follows from that of the parent MIMIC dataset, therefore, no additional ethical approval and consent were not required.

**Informed Consent:** Not applicable.

**Peer-review:** Externally peer-reviewed.

**Author Contributions:** Concept - L.L., J. Z., X.L.; Design - L.L., J.Z., X.L.; Supervision - L.L., J.Z., X.L.; Resources - L.L., J.Z., X.L.; Materials - L.L., J.Z., X.L.; Data Collection and/orProcessing - L.L., J.Z., X.L.; Analysis and/or Interpretation - L.L., J.Z., X.L.; Literature Search - L.L., J.Z., X.L.; Writing - L.L., J.Z., X.L., Critical Review - L.L., J.Z., X.L.

**Declaration of Interests:** The authors have no conflict of interest to declare.

**Funding:** The authors declared that this study has received no financial support.

## REFERENCES

1. Wong ND. Epidemiological studies of CHD and the evolution of preventive cardiology. *Nat Rev Cardiol*. 2014;11(5):276-289. [CrossRef]
2. Shahjehan RD, Sharma S, Bhutta BS. Coronary artery disease. *StatPearls*. 2025.
3. Ralapanawa U, Sivakanesan R. Epidemiology and the magnitude of coronary artery disease and acute coronary syndrome: a narrative review. *J Epidemiol Glob Health*. 2021;11(2):169-177. [CrossRef]
4. Kohsaka S, Ejiri K, Takagi H, et al. Diagnostic and therapeutic strategies for stable coronary artery disease following the ischemia trial. *JACC Asia*. 2023;3(1):15-30. [CrossRef]
5. Hetherington I, Totary-Jain H. Anti-atherosclerotic therapies: milestones, challenges, and emerging innovations. *Mol Ther*. 2022;30(10):3106-3117. [CrossRef]
6. Omeh DJ, Shlofmitz E. Restenosis of stented coronary arteries. *StatPearls*. 2025.

7. Montrief T, Koyfman A, Long B. Coronary artery bypass graft surgery complications: a review for emergency clinicians. *Am J Emerg Med*. 2018;36(12):2289-2297. [CrossRef]
8. Tantry US, Gurbel PA. Antiplatelet drug resistance and variability in response: the role of antiplatelet therapy monitoring. *Curr Pharm Des*. 2013;19(21):3795-3815. [CrossRef]
9. Mink S, Drexel H, Leiherer A, et al. Anti-SARS-CoV-2 antibodies versus Vaccination Status in CAD Patients with COVID-19: a Prospective, Propensity Score-Matched Cohort Study. *Vaccines (Basel)*. 2024;12(8):855. [CrossRef]
10. Ongel EA, Karakurt Z, Salturk C, et al. How do COPD comorbidities affect ICU outcomes? *Int J Chron Obstruct Pulmon Dis*. 2014;9:1187-1196. [CrossRef]
11. Lambden S, Laterre PF, Levy MM, Francois B. The SOFA score-development, utility and challenges of accurate assessment in clinical trials. *Crit Care*. 2019;23(1):374. [CrossRef]
12. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. Apache II: a severity of disease classification system. *Crit Care Med*. 1985;13(10):818-829. [CrossRef]
13. Pellathy TP, Pinsky MR, Hravnak M. Intensive Care Unit scoring systems. *Crit Care Nurse*. 2021;41(4):54-64. [CrossRef]
14. Reddy V, Reddy H, Gemnani R, Kumar S, Acharya S. Navigating the complexity of scoring systems in sepsis management: a comprehensive review. *Cureus*. 2024;16(2):e54030. [CrossRef]
15. Schwalbe N, Wahl B. Artificial intelligence and the future of global health. *Lancet*. 2020;395(10236):1579-1586. [CrossRef]
16. Zhao X, Lu Y, Li S, et al. Predicting renal function recovery and short-term reversibility among acute kidney injury patients in the ICU: comparison of machine learning methods and conventional regression. *Ren Fail*. 2022;44(1):1326-1337. [CrossRef]
17. Luo X, Li B, Zhu R, et al. Development and validation of an interpretable machine learning model for predicting in-hospital mortality for ischemic stroke patients in ICU. *Int J Med Inform*. 2025;198:105874. [CrossRef]
18. Johnson AEW, Bulgarelli L, Shen L, et al. MIMIC-IV, a freely accessible electronic health record dataset. *Sci Data*. 2023;10(1):1. [CrossRef]
19. Li X, Chen Y, Yuan Q, Zhou H, Lu L, Guo R. Neutrophil-to-lymphocyte ratio, monocyte-to-lymphocyte ratio, platelet-to-lymphocyte ratio associated with 28-day all-cause mortality in septic patients with coronary artery disease: a retrospective analysis of MIMIC-IV database. *BMC Infect Dis*. 2024;24(1):749. [CrossRef]
20. Si Y, Abdollahi A, Ashrafi N, et al. Optimized feature selection and advanced machine learning for stroke risk prediction in revascularized coronary artery disease patients. *BMC Med Inform Decis Mak*. 2025;25(1):276. [CrossRef]
21. Friedman JH, Meulman JJ. Multiple additive regression trees with application in epidemiology. *Stat Med*. 2003;22(9):1365-1381. [CrossRef]
22. Liu TD, Sun S, Zheng YY, Zhang JY. Aggregate index of systemic inflammation as a new prognostic marker in patients with coronary artery disease undergoing PCI. *Sci Rep*. 2025;15(1):37623. [CrossRef]
23. Wang J, Wang X, Yu W, Zhang K, Wei Y. Obstructive sleep apnea-induced multi-organ dysfunction after elective coronary artery bypass surgery in coronary heart disease patients. *J Thorac Dis*. 2020;12(10):5603-5616. [CrossRef]
24. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21(1):128-138. [CrossRef]
25. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26(6):565-574. [CrossRef]
26. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. 2015;162(1):55-63. [CrossRef]
27. Shawon MSR, Odutola M, Falster MO, Jorm LR. Patient and hospital factors associated with 30-day readmissions after coronary artery bypass graft (CABG) surgery: a systematic review and meta-analysis. *J Cardiothorac Surg*. 2021;16(1):172. [CrossRef]
28. Polsook R, Aungsuroch Y. A cross-sectional study of factors predicting readmission in Thais with coronary artery disease. *J Res Nurs*. 2021;26(4):293-304. [CrossRef]
29. Yang L, Wang Z, Zhao S, et al. Effectiveness of digital healthcare to improve clinical outcomes in discharged patients with coronary artery disease. *NPJ Digit Med*. 2025;8(1):473. [CrossRef]
30. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373-383. [CrossRef]
31. Cheraghi M, Karimi A, Sedighi M, Jabarzare S. The interplay between diabetes and atherosclerosis: a review of pathophysiological mechanisms. *ARYA Atheroscler*. 2025;21(5):62-80. [CrossRef]
32. Lim YJ, Sidor NA, Tonial NC, Che A, Urquhart BL. Uremic toxins in the progression of chronic kidney disease and cardiovascular disease: mechanisms and therapeutic targets. *Toxins (Basel)*. 2021;13(2):142. [CrossRef]
33. Coyle M, Flaherty G, Jennings C. A critical review of chronic kidney disease as a risk factor for coronary artery disease. *Int J Cardiol Heart Vasc*. 2021;35:100822. [CrossRef]
34. Palit S, Kendrick J. Vascular calcification in chronic kidney disease: role of disordered mineral metabolism. *Curr Pharm Des*. 2014;20(37):5829-5833. [CrossRef]
35. Efird JT, O'Neal WT, Davies SW, et al. Long-Term Mortality of 306,868 Patients with multi-vessel coronary artery disease: CABG versus Primate Conservation. *Br J Med Med Res*. 2013;3(4):1248-1257. [CrossRef]
36. Kirov H, Caldonazo T, Riedel LL, et al. Comparing outcomes between coronary artery bypass grafting and percutaneous coronary intervention in octogenarians with left main or multivessel disease. *Sci Rep*. 2023;13(1):22323. [CrossRef]
37. Kaimori JY, Sakaguchi Y, Kajimoto S, et al. Diagnosing metabolic acidosis in chronic kidney disease: importance of blood pH and serum anion gap. *Kidney Res Clin Pract*. 2022;41(3):288-297. [CrossRef]
38. Hosten AO. BUN and Creatinine. In: Walker HK, Hall WD, Hurst JW, editors. Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd ed. Boston: Butterworths; 1990. Chapter 193. PMID: 21250147.
39. Alani H, Tamimi A, Tamimi N. Cardiovascular co-morbidity in chronic kidney disease: current knowledge and future research needs. *World J Nephrol*. 2014;3(4):156-168. [CrossRef]

**Supplementary Table 1. Candidate variables**

| Variables | Descriptions | Categories |
|---|---|---|
| Demographics | | |
| age | Age | |
| sex | Sex | |
| race_simple5 | Classification of race/ethnicity | White, black, Asian, hispanic/ latino, and others |
| height | Height | |
| weight | Body weight | |
| insurance | Health insurance status | |
| disease history and prior treatment | | |
| cabg | History of coronary artery bypass grafting (CABG) | |
| pci | History of percutaneous coronary intervention (PCI) | |
| Organ Function Scores | | |
| Charlson comorbidity index | Charlson comorbidity index | |
| sofa | Sequential organ failure assessment (SOFA) score | |
| gcs_min | Minimum total Glasgow Coma Scale (GCS) score | |
| gcs_motor | GCS motor score | |
| gcs_verbal | GCS verbal score | |
| gcs_eyes | GCS eye score | |
| gcs_unable | GCS cannot be assessed | |
| Vital Signs | | |
| heart_rate_max | Maximum heart rate | |
| sbp_min | Minimum systolic blood pressure | |
| dbp_min | Minimum diastolic blood pressure | |
| mbp_min | Minimum mean arterial pressure | |
| resp_rate_max | Maximum respiratory rate | |
| temperature_max | Maximum body temperature | |
| spo2_min | Minimum pulse oxygen saturation ($SpO_2$) | |
| glucose_max | Maximum blood glucose level | |
| Laboratory Values | | |
| hematocrit_min | Minimum hematocrit (HCT) | |
| hematocrit_max | Maximum hematocrit (HCT) | |
| hemoglobin_min | Minimum hemoglobin (Hb) | |
| platelets_min | Minimum platelet count | |
| wbc_max | Maximum white blood cell count (WBC) | |
| aniongap_max | Maximum anion gap | |
| bicarbonate_min | Minimum bicarbonate ($HCO_3^-$) | |
| bun_max | Maximum blood urea nitrogen (BUN) | |
| calcium_min | Minimum serum calcium ($Ca^{2+}$) | |
| chloride_max | Maximum chloride ($Cl^-$) | |
| creatinine_max | Maximum serum creatinine (Cr) | |
| sodium_min | Minimum sodium ($Na^+$) | |
| potassium_max | Maximum potassium ($K^+$) | |
| abs_basophils_max | Maximum absolute basophil count | |
| abs_eosinophils_max | Maximum absolute eosinophil count | |
| abs_lymphocytes_min | Minimum absolute lymphocyte count | |
| abs_monocytes_min | Minimum absolute monocyte count | |
| abs_neutrophils_max | Maximum absolute neutrophil count | |
| inr_max | Maximum International Normalized Ratio (INR) | |

| Variables | Descriptions | Categories |
|---|---|---|
| pt_max | Maximum prothrombin time (PT) | |
| ptt_max | Maximum activated partial thromboplastin time (aPTT) | |
| Electrocardiogram Parameters | | |
| rr_interval | RR interval | |
| QRS_duration | QRS duration | |
| QT_interval | QT interval | |
| p_axis | P wave axis | |
| qrs_axis | QRS axis | |
| t_axis | T wave axis | |
| Composite Measure | | |
| NLR | Neutrophil-to-lymphocyte ratio | Derived from calculation |
| MLR | Monocyte-to-lymphocyte ratio | Derived from calculation |

**Supplementary Table 2. Comparison of baseline characteristics of key variables**

| Variable | Total | Train (3,830) | Test (1,642) | *P*_value |
|---|---|---|---|---|
| CABG | 1860 (34%) | 1301 (34%) | 559 (34%) | 0.9819 |
| Charlson comorbidity index | 4.91 (2.57) | 4.9 (2.57) | 4.94 (2.57) | 0.6353 |
| Aniongap | 14.95 (4.45) | 14.85 (4.38) | 15.13 (4.61) | 0.0391 |
| SOFA | 4.74 (3.06) | 4.69 (3.05) | 4.85 (3.06) | 0.0794 |
| NLR | 7.6 (8.47) | 7.17 (7.59) | 7.01 (6.84) | 0.4479 |
| Spo2 | 92.33 (5.18) | 92.37 (4.93) | 92.22 (5.71) | 0.3719 |
| Respiratory rate | 27.32 (5.66) | 27.37 (5.65) | 27.23 (5.71) | 0.4008 |
| Weight | 86.42 (20.34) | 86.08 (20.16) | 86.97 (20.72) | 0.1441 |
| PT | 15.91 (7.3) | 15.87 (7) | 15.89 (7.73) | 0.9522 |
| BUN | 24.08 (17.32) | 23.99 (17.15) | 24.3 (17.75) | 0.5542 |
| PTT | 31.6 (12.38) | 31.55 (11.95) | 31.88 (13.33) | 0.3985 |
| QRS_duration | 102.28 (24.7) | 102.45 (24.91) | 101.85 (24.17) | 0.3972 |
| Heart rate | 98.6 (16.99) | 98.69 (17.39) | 98.4 (16) | 0.5483 |
| Died_1year | 777 (14.2%) | 544 (14.2%) | 233 (14.2%) | 1 |
| Died_28days_inhospital | 304 (5.6%) | 210 (5.5%) | 94 (5.7%) | 0.7693 |
| Has_readmit | 496 (9.1%) | 319 (8.3%) | 177 (10.8%) | 0.0045 |

Abbreviations: SBP, systolic blood pressure; DBP, diastolic blood pressure; MBP, mean blood pressure; GCS, Glasgow Coma Scale; NLR, neutrophil-to-lymphocyte ratio; MLR, monocyte-to-lymphocyte ratio; ICU, intensive care unit; PT, prothrombin time; PTT, partial thromboplastin time; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; Crp, C-reactive protein.

**Supplementary Table 3. Comparison of patients with available versus missing ECG data**

| Variable | ECG available (n=5,472) | ECG missing (n=1,305) | SMD |
|---|---|---|---|
| Age | 67.69 (11.92) | 68.04 (11.56) | -0.030 |
| Male | 3925/5472 (71.7%) | 936/1305 (71.7%) | 0.000 |
| Charlson comorbidity index | 4.91 (2.57) | 5.55 (2.81) | -0.238 |
| SOFA score | 4.74 (3.06) | 5.03 (3.03) | -0.096 |
| 1-year all-cause mortality | 777/5472 (14.2%) | 247/1305 (18.9%) | -0.127 |
| 28-day in-hospital mortality | 304/5472 (5.6%) | 104/1305 (8.0%) | -0.096 |
| CAD-related ICU readmission | 496/5472 (9.1%) | 129/1305 (9.9%) | -0.028 |

Continuous variables are presented as mean (SD); categorical variables as n/N (%). SMD, standardized mean difference. Positive SMD indicates a higher value in the ECG-available group.

**Supplementary Table 4. Prespecified model parameters and outcome-specific classification thresholds**

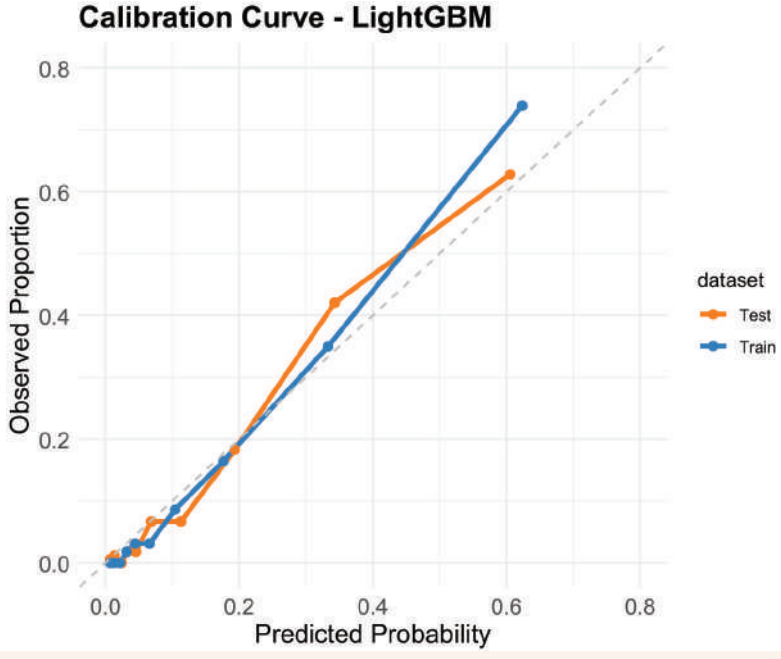
| Model | Common parameters | 1-year mortality | 28-day mortality | CAD-related ICU readmission |
|---|---|---|---|---|
| Logistic regression | Maximum-likelihood estimation; no tunable hyperparameters | Threshold=0.18 | Threshold=0.06 | Threshold=0.08 |
| Random forest | ntree=300; maxnodes=10 | Threshold=0.18 | Threshold=0.06 | Threshold=0.08 |
| Support vector machine | Radial kernel; cost=1; gamma=1/13; probability=TRUE | Threshold=0.18 | Threshold=0.06 | Threshold=0.08 |
| XGBoost | nrounds=20; max_depth=3; eta=0.1; objective=binary:logistic | Threshold=0.18 | Threshold=0.06 | Threshold=0.08 |
| AdaBoost | iter=20; nu=1; type=discrete | Base-tree maxdepth=3; threshold=0.18 | Base-tree maxdepth=3; threshold=0.06 | Base-tree maxdepth=2; threshold=0.08 |
| Naive Bayes | laplace=0; package defaults for continuous predictors | Threshold=0.18 | Threshold=0.06 | Threshold=0.08 |
| K-nearest neighbors | Predictors standardized using training-set mean and SD | k=20; threshold=0.18 | k=50; threshold=0.06 | k=50; threshold=0.08 |
| CatBoost | iterations=20; learning_rate=0.1; loss=Logloss | depth=3; threshold=0.18 | depth=3; threshold=0.06 | depth=2; threshold=0.08 |
| LightGBM | learning_rate=0.1; num_leaves=20; objective=binary | nrounds=50; max_depth=3; threshold=0.18 | nrounds=100; max_depth=2; threshold=0.06 | nrounds=100; max_depth=2; threshold=0.08 |

Parameters were prespecified before evaluation in the held-out test set. The test set was not used for hyperparameter tuning or threshold optimization. Outcome-specific thresholds were fixed at 0.18, 0.06, and 0.08, respectively, and applied uniformly across models.

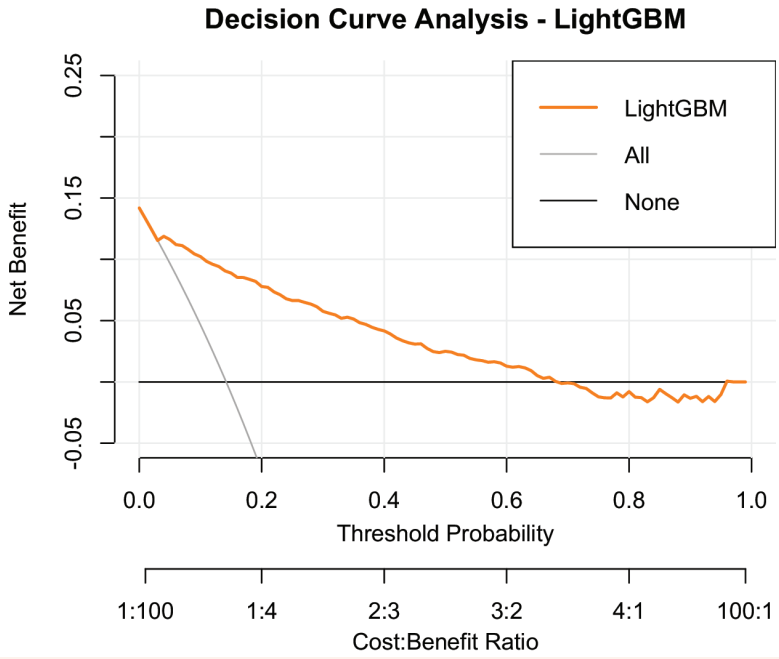
**Supplementary Table 5 Quantitative calibration of LightGBM in the held-out test set**

| Outcome | Brier score | Calibration intercept (95% CI) | Calibration slope (95% CI) |
|---|---|---|---|
| 1-year all-cause mortality | 0.081 | 0.170 (-0.057, 0.397) | 1.179 (1.040, 1.318) |
| 28-day in-hospital mortality | 0.043 | 0.098 (-0.296, 0.493) | 1.032 (0.865, 1.198) |
| CAD-related ICU readmission | 0.094 | -0.203 (-0.675, 0.269) | 0.777 (0.582, 0.973) |

Ideal calibration corresponds to an intercept of 0 and a slope of 1. A slope below 1 suggests overly extreme predictions and potential overfitting.




**Supplementary Figure 1. Calibration curve of LightGBM Model for one-year mortality prediction.**

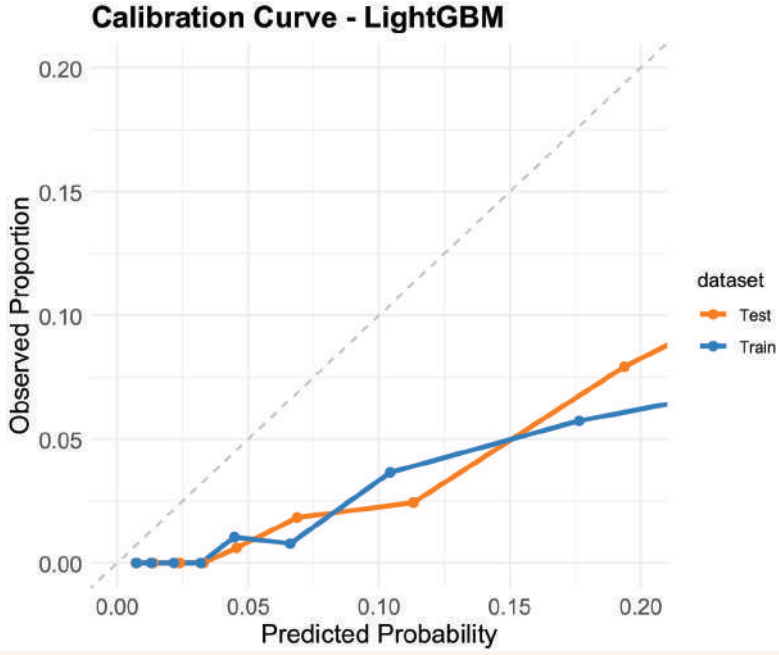



**Supplementary Figure 2. Decision curve analysis of LightGBM Model for one-year mortality prediction.**

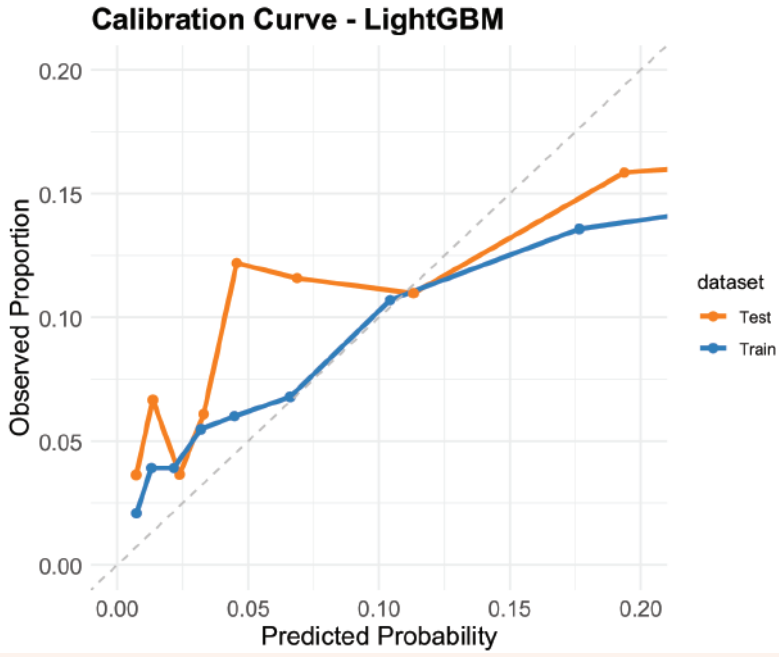
**Supplementary Table 6. Paired DeLong comparison of LightGBM and logistic regression in the held-out test set**

| Outcome | LightGBM AUC | Logistic regression AUC | AUC difference (95% CI) | P value |
|---|---|---|---|---|
| 1-year all-cause mortality | 0.896 | 0.887 | 0.0091 (-0.0004, 0.0186) | 0.062 |
| 28-day in-hospital mortality | 0.898 | 0.894 | 0.0036 (-0.0117, 0.0190) | 0.642 |
| CAD-related ICU readmission | 0.690 | 0.673 | 0.0166 (-0.0110, 0.0442) | 0.238 |

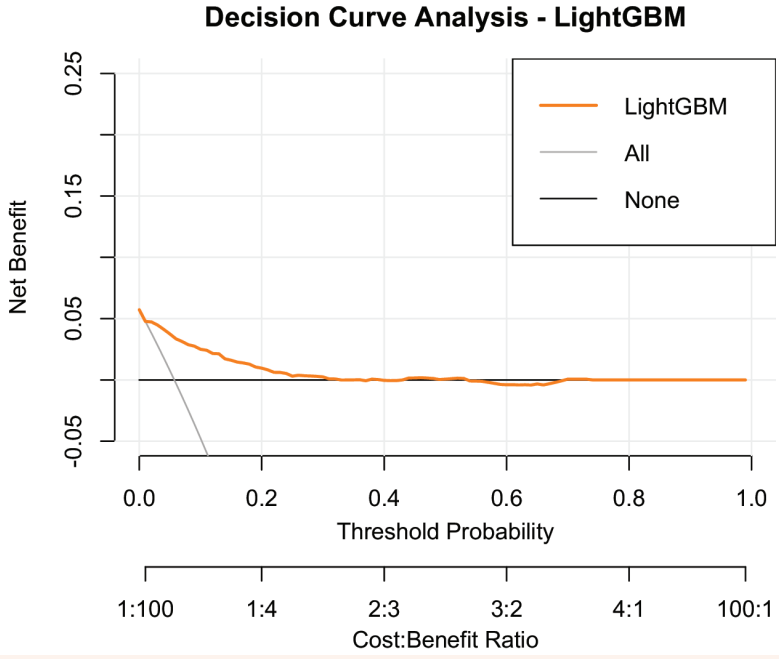
AUC difference was calculated as LightGBM minus logistic regression. All P values are two-sided. None of the three comparisons reached P<0.05.



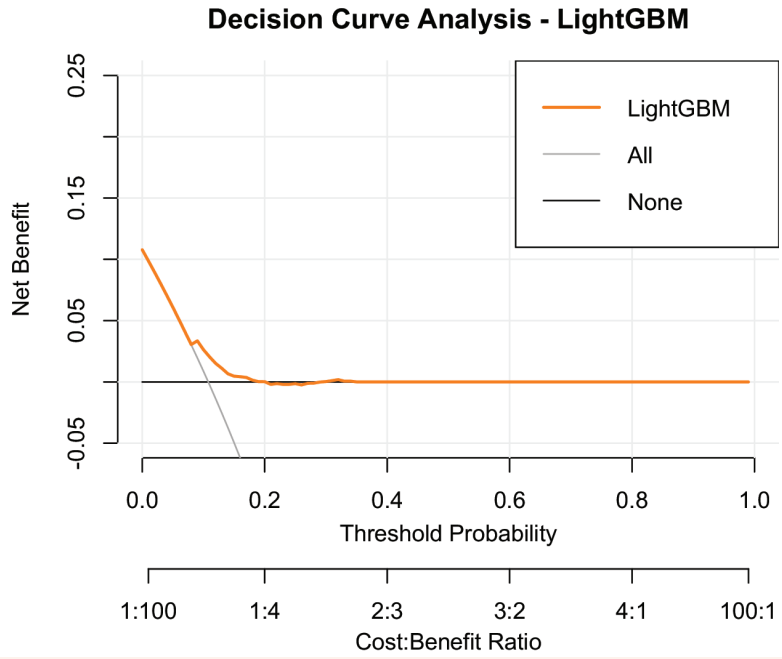

**Supplementary Figure 3. Calibration curve of the LightGBM model for 28-day in-hospital mortality prediction.**




**Supplementary Figure 4. Calibration curve of the LightGBM model for CAD-related ICU readmission prediction.**




**Supplementary Figure 5. Decision curve analysis of the LightGBM model for 28-day in-hospital mortality prediction.**

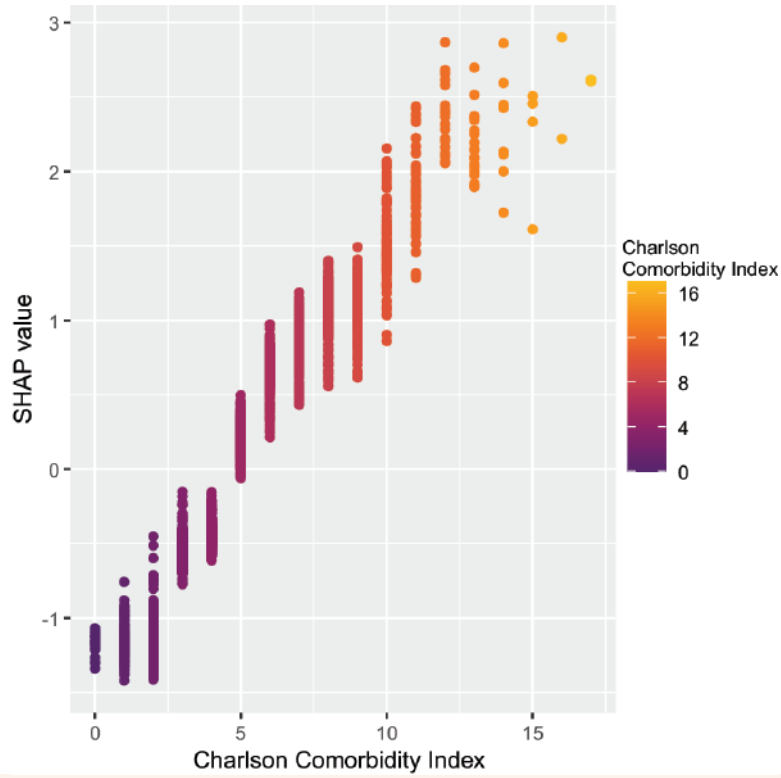



**Supplementary Figure 6. Decision curve analysis of the LightGBM model for CAD-related ICU readmission prediction.**

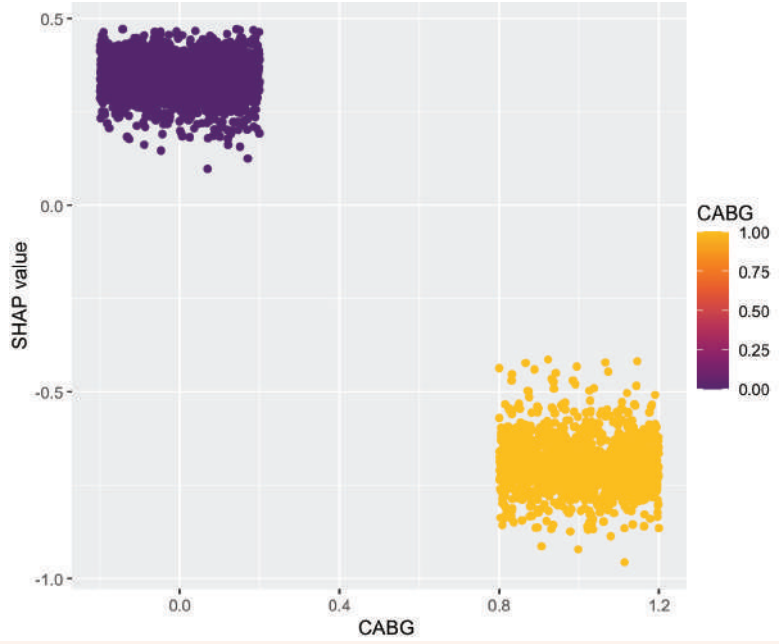
**Supplementary Table 7. Five-imputation sensitivity analysis in the held-out test set**

| Outcome | Model | Main analysis AUC (95% CI) | Five-imputation AUC (95% CI) | AUC change |
|---|---|---|---|---|
| 1-year all-cause mortality | LightGBM | 0.896 (0.875, 0.917) | 0.897 (0.876, 0.918) | +0.001 |
| 1-year all-cause mortality | Logistic Regression | 0.887 (0.864, 0.911) | 0.887 (0.863, 0.910) | -0.000 |
| 28-day in-hospital mortality | LightGBM | 0.898 (0.870, 0.925) | 0.898 (0.870, 0.926) | +0.000 |
| 28-day in-hospital mortality | Logistic Regression | 0.894 (0.862, 0.926) | 0.896 (0.863, 0.928) | +0.001 |
| CAD-related ICU readmission | LightGBM | 0.690 (0.651, 0.728) | 0.684 (0.645, 0.722) | -0.006 |
| CAD-related ICU readmission | Logistic Regression | 0.673 (0.633, 0.713) | 0.665 (0.625, 0.705) | -0.008 |

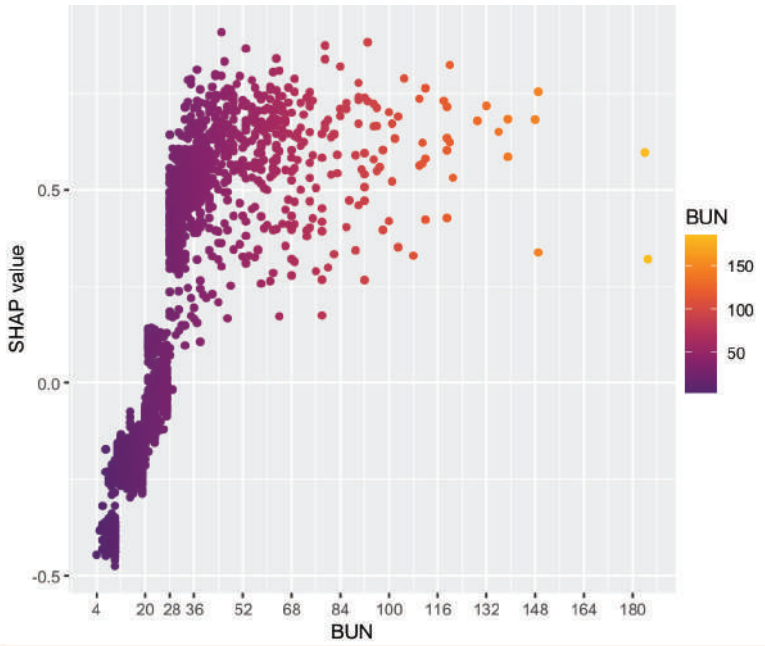
Five imputation models were estimated without using test-set rows; each completed training set was modeled independently, and participant-level probabilities were averaged across the five fitted models. The ranking and interpretation were unchanged.



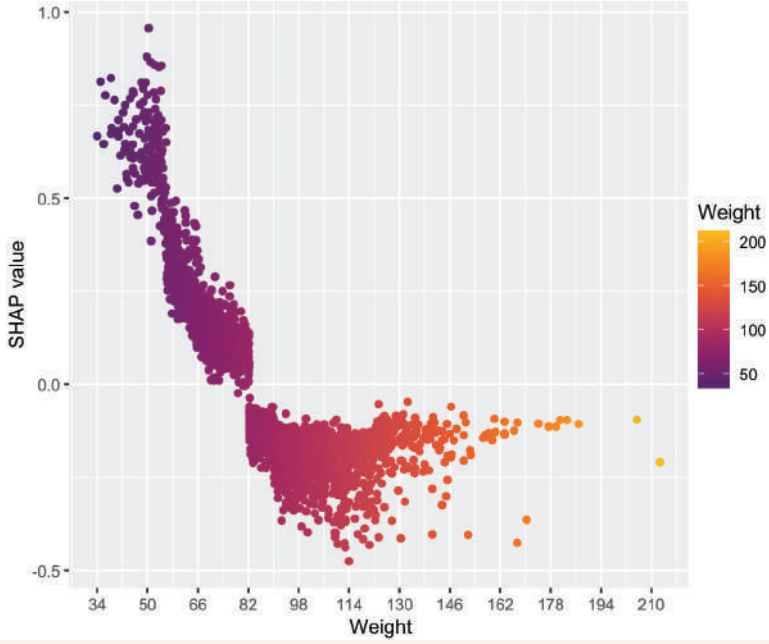

**Supplementary Figure 7. SHAP dependence plot of Charlson comorbidity index for one-year mortality prediction.**



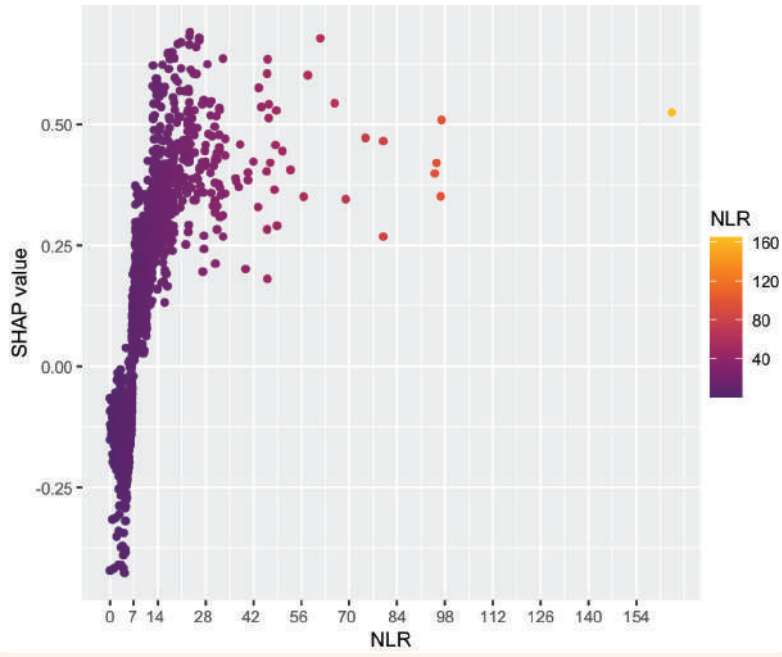

**Supplementary Figure 8. SHAP dependence plot of CABG for one-year mortality prediction.**



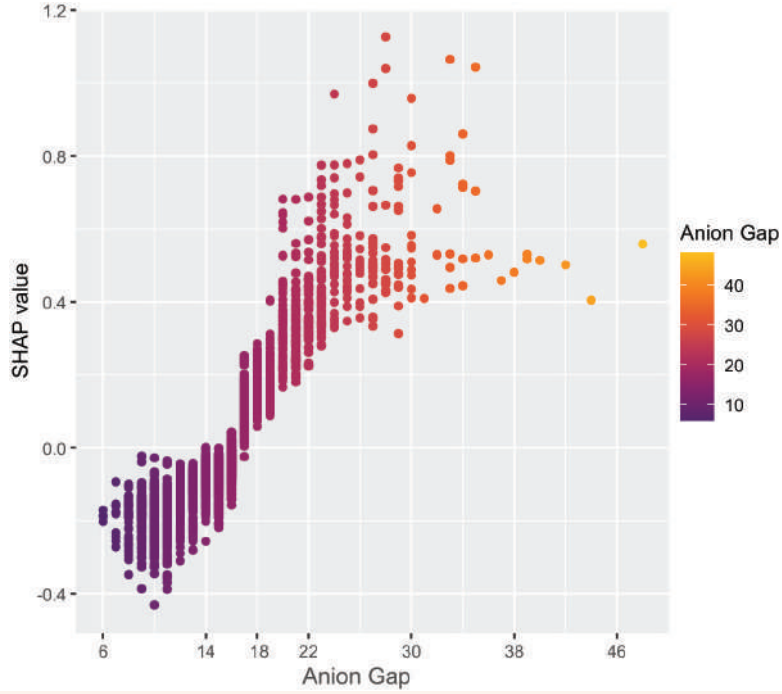

**Supplementary Figure 9. SHAP dependence plot of BUN for one-year mortality prediction.**



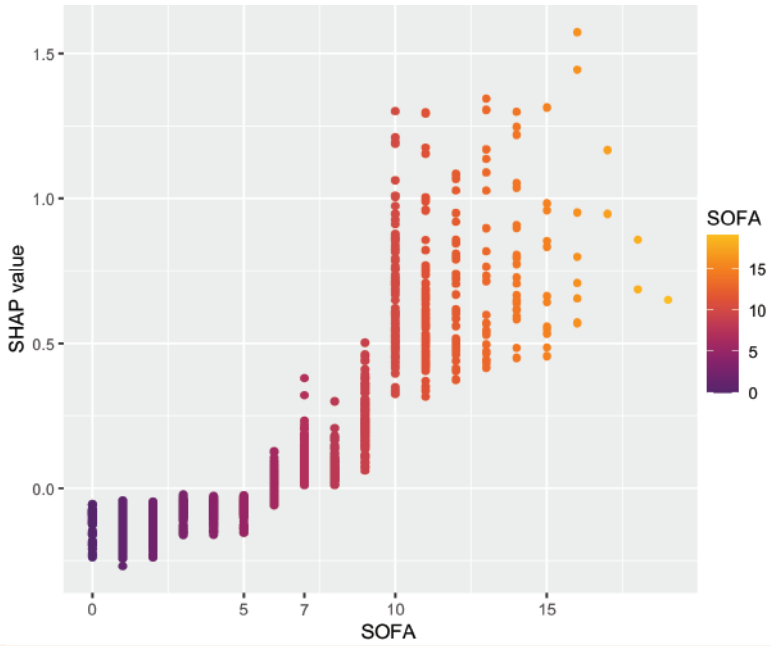

**Supplementary Figure 10. SHAP dependence plot of weight for one-year mortality prediction.**



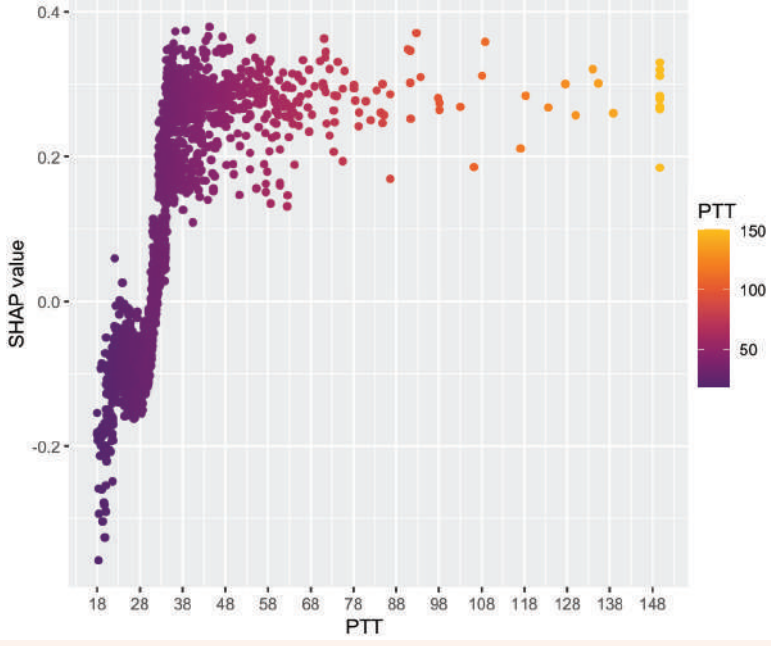

**Supplementary Figure 11. SHAP dependence plot of NLR for one-year mortality prediction.**



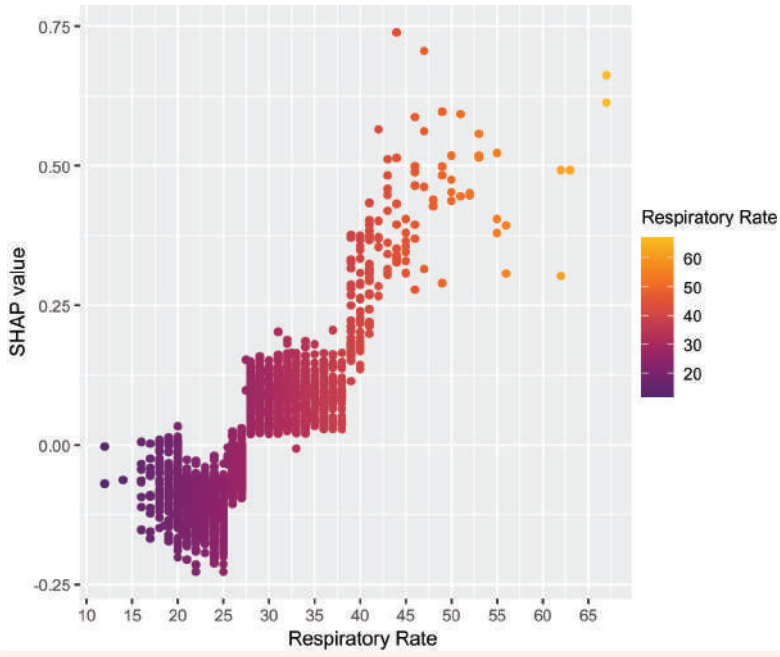

**Supplementary Figure 12. SHAP dependence plot of anion gap for one-year mortality prediction.**



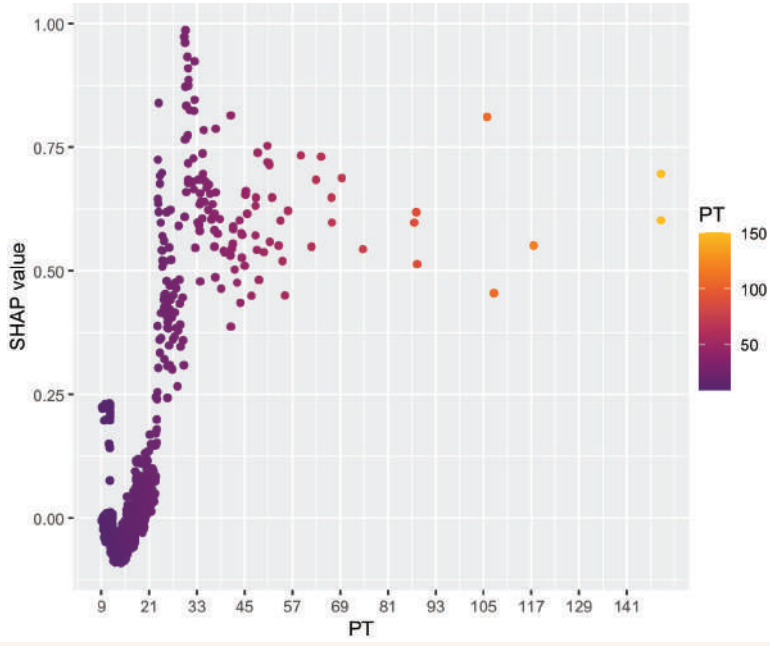

**Supplementary Figure 13. SHAP dependence plot of SOFA for one-year mortality prediction.**



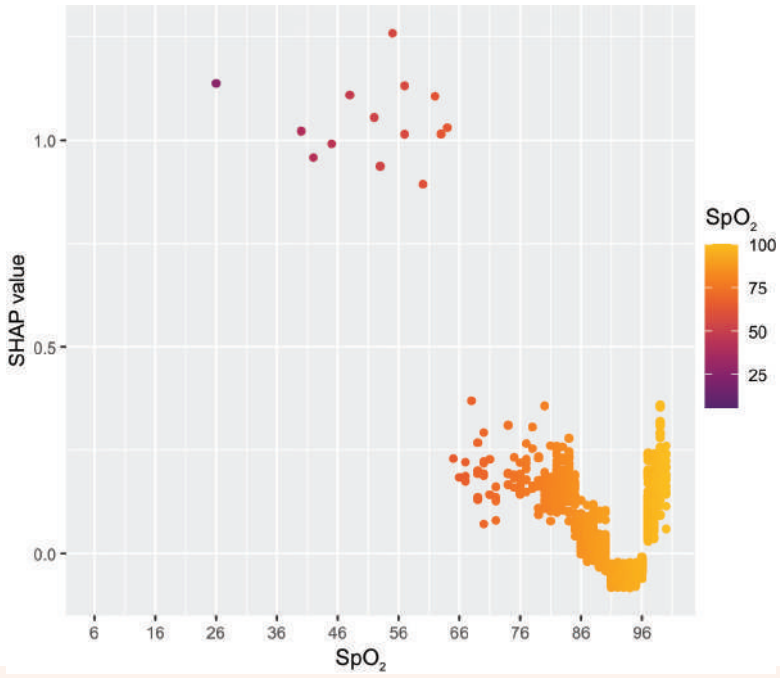

**Supplementary Figure 14. SHAP dependence plot of PTT for one-year mortality prediction.**



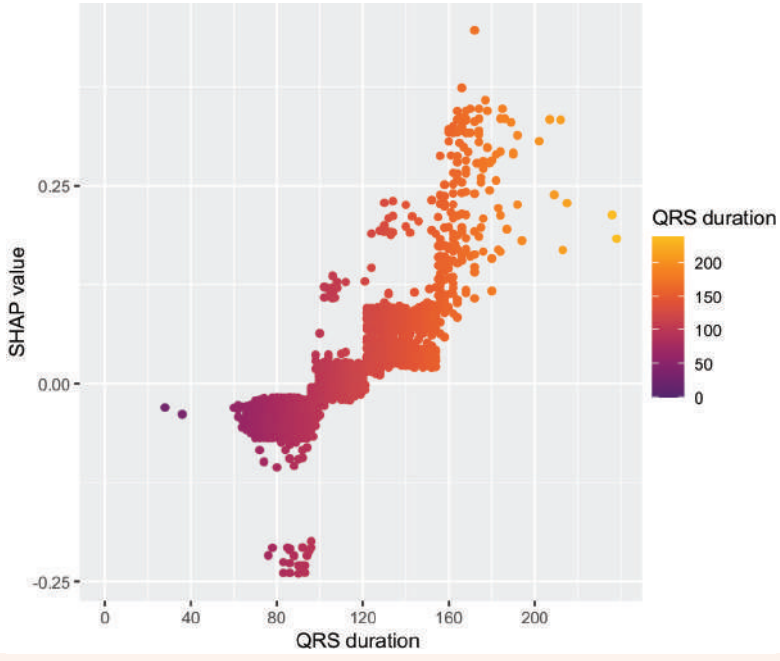

**Supplementary Figure 15. SHAP dependence plot of respiratory rate for one-year mortality prediction.**



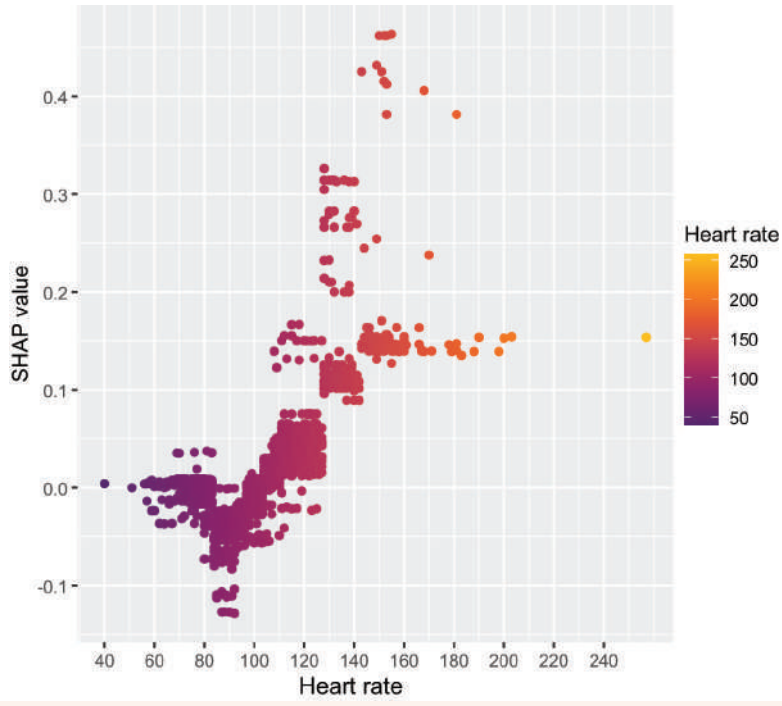

**Supplementary Figure 16. SHAP dependence plot of PT for one-year mortality prediction.**



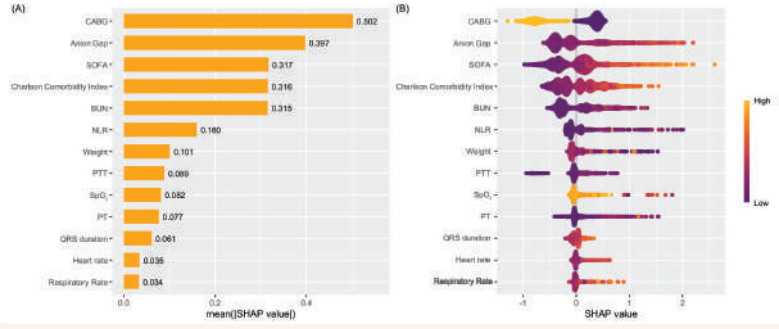

**Supplementary Figure 17. SHAP dependence plot of $SpO_2$ for one-year mortality prediction.**



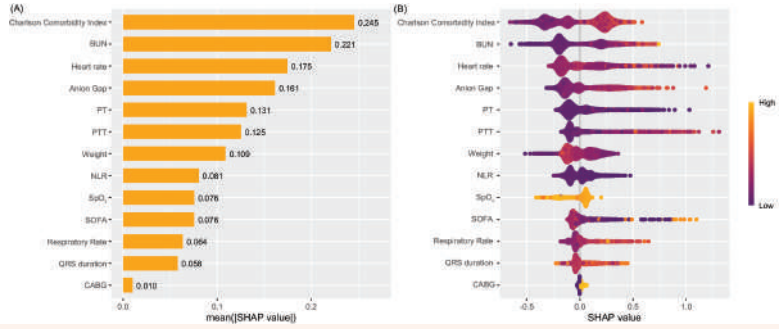

**Supplementary Figure 18. SHAP dependence plot of QRS duration for one-year mortality prediction.**




**Supplementary Figure 19. SHAP dependence plot of heart rate for one-year mortality prediction.**




**Supplementary Figure 20. SHAP feature importance and summary for 28-day in-hospital mortality prediction.**




**Supplementary Figure 21. SHAP feature importance and summary for CAD-related ICU readmission prediction.**