Research Square

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

Development and external validation of an interpretable machine learning model for in-hospital major adverse cardiovascular events after PCI in acute coronary syndrome

Dongdong Yan

yandd8888@163.com

Lanzhou University

Erkun Xin

Lanzhou University

Yongxiang Wang

Lanzhou University

Wei Tian

Gansu University of Traditional Chinese Medicine

Wenqiang Li

Qinghai Cardio Cerebrovascular Specialty Hospital

Tao Li

Gansu Gem Flower Hospital

Zheng Zhang

Lanzhou University

Research Article

Keywords: acute coronary syndrome, machine learning, electronic medical records, major adverse cardiovascular events, percutaneous coronary intervention, random forest, external validation, clinical decision support

Posted Date: November 7th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11290669/v1

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

Additional Declarations: No competing interests reported.

Abstract

Background

Electronic medical record (EMR) data provide an opportunity to support early risk stratification in acute coronary syndrome (ACS), but models with large predictor sets may be difficult to implement. We aimed to develop and externally validate a parsimonious, interpretable machine learning model for in-hospital major adverse cardiovascular events (MACE) in patients with ACS undergoing percutaneous coronary intervention (PCI). Methods: This two-center retrospective cohort study included 1,726 patients: 1,218 in the derivation cohort and 508 in an independent external validation cohort. The primary outcome was in-hospital MACE, defined as a composite of death, cerebral hemorrhage, cerebral infarction, cardiogenic shock, cardiac rupture, recurrent myocardial infarction, stent thrombosis, or Bleeding Academic Research Consortium bleeding events. A total of 125 candidate EMR variables were evaluated. LASSO and Boruta-based feature selection, alternative preprocessing strategies including ADASYN, seven machine learning algorithms, and grid-search cross-validation were compared. Discrimination was assessed using AUROC and AUPRC, with calibration and decision-curve analyses used to assess agreement and potential clinical utility. Results: In-hospital MACE occurred in 138 of 1,218 patients (11.3%) in the derivation cohort and 15 of 508 patients (3.0%) in the external cohort. The final random forest model retained eight predictors: Killip class, inorganic phosphorus, extracorporeal membrane oxygenation, left ventricular ejection fraction, absolute neutrophil count, total protein, alanine aminotransferase, and potassium. The model achieved an AUROC of 0.82 (95% CI 0.73–0.90) in internal validation and 0.90 (95% CI 0.85–0.95) in external validation. At the reported operating point, sensitivity was 0.488 and specificity was 0.991. Conclusions: An eight-variable random forest model derived from routinely collected EMR data showed good discrimination and preserved ranking performance in an independent hospital cohort. The model is best considered an early in-hospital risk-stratification tool; prospective validation with a prespecified prediction time is required before clinical implementation.

Introduction

Acute coronary syndrome (ACS) represents an unstable state of coronary artery disease (CAD) and encompasses a group of life-threatening conditions that frequently require emergent intervention to improve prognosis, with varying degrees of urgency across subtypes[1, 2, 3]. To facilitate timely identification of these critical illnesses, numerous prediction models have been developed for distinct disease subtypes[4, 5]. The widespread establishment of chest pain centers has substantially reduced in-hospital mortality, paradoxically increasing the difficulty of model development due to lower event rates[6, 7]. The advent of machine learning (ML) has enhanced the performance of these predictive models to varying extents, with some achieving individualized risk stratification[8, 9]. However, feature selection methods remain relatively homogeneous across studies, and the number of included predictors varies considerably[10, 11].

Patients with ACS remain at risk of serious in-hospital complications despite contemporary PCI. Established risk scores such as GRACE and newer machine-learning approaches have improved prognostic assessment, but many models use fixed or relatively large predictor sets and may be difficult to transport across institutions[12, 13, 14, 15]. Prediction studies also frequently focus on discrimination alone, while calibration, clinical utility, class imbalance, model parsimony, and external validation receive less attention[4, 16, 17, 18]. Prior work has examined risk prediction in specialized ACS populations, including patients with cancer [19]. We therefore

selected in-hospital MACE as a clinically relevant composite outcome and evaluated whether routinely available EMR variables could support a compact risk-stratification model. Rather than assuming that synthetic oversampling or a particular algorithm would be beneficial, we compared alternative preprocessing and modeling strategies and assessed their performance using both AUROC and AUPRC.

The objectives of this study were to: (1) derive a parsimonious prediction model from a broad set of routinely collected EMR variables; (2) compare sequential feature-selection and class-imbalance strategies; (3) compare seven machine-learning algorithms using consistent internal validation; (4) evaluate model interpretation using feature-importance, LIME, and SHAP analyses; and (5) examine transportability in an independent hospital cohort. The study is reported with reference to the principles of transparent reporting for multivariable prediction-model research.

Materials and methods

Study Population

A total of 1,218 patients with ACS who underwent emergency PCI at The First Hospital of Lanzhou University between January and December 2024 were included in the model development cohort. In contrast, 508 ACS patients treated at Gansu Gem Flower Hospital from January 2024 to December 2025 were included in the external validation cohort (Fig. 1). The diagnosis of ACS was established in accordance with the 2025 American Heart Association guidelines for the management of patients with ACS. The inclusion criteria were as follows: (1) a confirmed diagnosis of STEMI or NSTE-ACS, (2) age ≥ 18 years, (3) receipt of PCI, and (4) availability of complete clinical data. Detailed inclusion and exclusion processes are illustrated in Fig. 1.

Data collection

All patient data were collected through the EMRS of the cardiology center at the First Hospital of Lanzhou University and Gansu Gem Flower Hospital. Trained personnel were responsible for data entry, with dedicated staff overseeing data quality control and verification.

The collected variables encompassed demographic characteristics, vital signs, laboratory parameters, echocardiographic measurements, medications, and in-hospital outcomes. Demographic and clinical variables included sex, age, height, weight, body mass index (BMI), systolic blood pressure (SBP), diastolic blood pressure (DBP), heart rate (HR), ischemia time (IT), duration of intra-aortic balloon pump (IABP) support, use of extracorporeal membrane oxygenation (ECMO), pacemaker implantation, and past medical history of hypertension (HTN), coronary artery disease (CAD), type 2 diabetes mellitus (T2DM), stroke, atrial fibrillation (AF), chronic obstructive pulmonary disease (COPD), smoking history (SH), drinking history (DH), and family history of CAD (CADH). Coronary lesion locations were also recorded, including the left main (LM) artery, left anterior descending (LAD) artery, and left circumflex (LCX) artery.

All laboratory parameters were derived from the first blood sample collected after admission, prior to primary percutaneous coronary intervention (PCI). These included routine blood tests: white blood cell count (WBC), red blood cell count (RBC), hemoglobin (HB), hematocrit (HCT), platelet count (PLT), mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), red cell distribution width (RDW_CV and RDW_SD), platelet distribution width (PDW), mean platelet volume (MPV),

platelet crit (PCT), and platelet-large cell ratio (PLCR); differential leukocyte counts: neutrophil percentage (NEUT_P) and absolute count (NEUT_AC), lymphocyte percentage (LY_P) and absolute count (LY_AC), monocyte percentage (MO_P) and absolute count (MO_AC), eosinophil percentage (EO_P) and absolute count (EO_AC), basophil percentage (BA_P) and absolute count (BA_AC); reticulocyte parameters: reticulocyte percentage (RET_P), reticulocyte count (RET_C), immature reticulocyte fraction (IRF), low-fluorescence reticulocyte ratio (LFR), and medium-fluorescence reticulocyte ratio (MFR); immature granulocyte percentage (IG_P) and absolute count (IG); nucleated red blood cell percentage (NRBC_P) and count (NRBC_C); and coagulation-related markers: D-dimer (DDimer).

Biochemical parameters comprised: liver and kidney function tests—alanine aminotransferase (ALT), aspartate aminotransferase (AST), AST/ALT ratio, gamma-glutamyl transferase (GGT), alkaline phosphatase (ALP), cholinesterase (CHE), total protein (TP), albumin (ALB), globulin (GLB), albumin-to-globulin ratio (AGR), total bilirubin (TBIL), direct bilirubin (DBIL), indirect bilirubin (IBIL), total bile acid (TBA), urea, creatinine (Crea), urea-to-creatinine ratio (UCR), uric acid (UA), and α-L-fucosidase (AFU); lipid profile—total cholesterol (TC), triglyceride (TG), low-density lipoprotein cholesterol (LDLC), high-density lipoprotein cholesterol (HDLC), apolipoprotein A1 (ApoA1), apolipoprotein B (ApoB), apolipoprotein B/A1 ratio (ApoBAI), and lipoprotein(a) (Lp(a)); glycemic parameters—glucose (Glu), 2-hour plasma glucose (2h_PG); electrolytes—sodium (Na), potassium (K), calcium (Ca), magnesium (Mg), inorganic phosphorus (IP), chloride (Cl), and anion gap (AG); iron metabolism markers—iron (Fe), ferritin (SF), unsaturated iron-binding capacity (UIBC), total iron-binding capacity (TIBC), iron saturation (IS); and other biomarkers—N-terminal pro-B-type natriuretic peptide (NT-proBNP), high-sensitivity C-reactive protein (hsCRP), cardiac troponin I (cTnI), myoglobin (MYO), creatine kinase (CK), creatine kinase-MB (CK-MB), lactate dehydrogenase (LDH), α-hydroxybutyrate dehydrogenase (αHBDH), homocysteine (HCY), total IgE, thyroid function markers (thyroid-stimulating hormone [TSH], free triiodothyronine [FT3], free thyroxine [FT4], total triiodothyronine [TT3], total thyroxine [TT4], thyroglobulin [TGB], thyroglobulin antibody [TgAb], thyroid peroxidase antibody [TPOAb]), vitamin B12 (VB12), folate (FOL), erythropoietin (EPO), prealbumin (PA), urinary electrolytes (urine sodium [U_Na], urine potassium [U_K], urine calcium [U_Ca], urine phosphorus [U_P], urine chloride [U_Cl]), amylase (AMY), osmolality (OSM), serum index (SI), hemolysis index (HI), lipemia index (LI), and carbon dioxide (CO_2).

Echocardiographic parameters were based on the first ultrasound report obtained within 24 hours of admission, including left atrial diameter (LAD), mitral inflow peak E velocity (MV_E), mitral inflow peak A velocity (MV_A), left ventricular ejection fraction (LVEF), left ventricular end-diastolic volume (LVEDV), and left ventricular end-systolic volume (LVESV).

Medication history at admission was also recorded, encompassing antiplatelet agents (ticagrelor [TICA]), statins (atorvastatin [ATV], rosuvastatin [RSV]), fibrates (FIB), ezetimibe (EZT), ACE inhibitors (ACEI), angiotensin receptor blockers (ARB), angiotensin receptor–neprilysin inhibitors (ARNI), beta-blockers (BB), calcium channel blockers (CCB), diuretics (DRT), and SGLT2 inhibitors (SGLT2i).

The primary endpoint of this study was in-hospital MACE, comprising death, cerebral hemorrhage, cerebral infarction, cardiogenic shock, cardiac rupture, recurrent myocardial infarction, stent thrombosis, and BARC bleeding events, with corresponding dates recorded for each event.

Model development and comparison

The derivation cohort was randomly divided into a training set (70%) and an internal validation set (30%). Missing values were handled using multiple imputation by chained equations (MICE). We evaluated alternative preprocessing pipelines rather than prespecifying synthetic oversampling as mandatory. These pipelines included adaptive synthetic sampling (ADASYN), LASSO feature selection with L1 regularization, BorutaShap feature selection, and combinations of these approaches with grid-search cross-validation (GSCV). LASSO with fivefold cross-validation was used to shrink non-informative coefficients toward zero, followed by Boruta-based refinement of the candidate feature set. Seven algorithms were evaluated: random forest (RF), LightGBM, support vector machine (SVM), XGBoost, logistic regression (LR), k-nearest neighbors (KNN), and deep neural network (DNN). Hyperparameters were tuned by GSCV. Model discrimination was compared using AUROC and AUPRC; accuracy, precision, sensitivity, specificity, and F1 score were reported as complementary threshold-dependent measures. The final pipeline was selected by considering predictive performance, parsimony, and clinical interpretability. Model interpretation included global feature importance, local interpretable model-agnostic explanations (LIME), and Shapley additive explanations (SHAP).

Statistical analysis

All statistical analyses and predictive modeling were performed using Python version 3.11.9 (Python Software Foundation, Wilmington, DE, USA). Continuous variables were summarized as mean ± standard deviation when approximately normally distributed and as median (Q1, Q3) otherwise; groups were compared using the independent-samples t test or Mann-Whitney U test, as appropriate. Categorical variables were summarized as counts (percentages) and compared using the chi-square test or Fisher exact test when expected cell counts were small. Missing data were imputed using MICE. Model discrimination was evaluated with AUROC and AUPRC. Bootstrap resampling (1,000 replications) was used for uncertainty estimation and pairwise performance comparisons. Calibration plots were used to compare predicted and observed risk, and decision-curve analysis was used to estimate net benefit across clinically relevant threshold probabilities. All tests were two-sided, and $P < 0.05$ was considered statistically significant.

Results

Population Characteristics

The derivation cohort comprised 1,218 patients with ACS treated at the First Hospital of Lanzhou University; 138 patients (11.3%) experienced at least one in-hospital MACE. Missingness was 9.3% or less for the listed variables and below 0.5% for most remaining variables; MICE was used to address missing data. The independent external validation cohort comprised 508 patients from Gansu Gem Flower Hospital, of whom 15 (3.0%) experienced in-hospital MACE. The lower event rate and marked differences in several baseline characteristics demonstrate substantial case-mix heterogeneity between the two centers (Table 1).

Table 1
Baseline characteristics of the derivation and external validation cohorts

Variables	Derivation(n = 1218)	External validation(n = 508)	*p* value
ECMO	0(0, 0)	0(0, 0)	< 0.001
TP	67.7(64, 71.5)	62(58, 67)	< 0.001
ALT	59.73 ± 248.05	44.48 ± 182.18	0.211
IP	0.95(0.8, 1.12)	1.02(0.97, 1.13)	< 0.001
NEUT_AC	7.66(5.74, 10.37)	5.57(4.01, 8.25)	< 0.001
K	3.76(3.52, 4.02)	4(3.68,4.3)	< 0.001
LVEF (%)	52.39 ± 7.46	57.16 ± 7.80	< 0.001
Killip1 (%)	995(81.69)	262(51.57)	< 0.001
Killip2 (%)	125(10.26)	154(30.31)	
Killip3 (%)	26(2.13)	44(8.66)	
Killip4 (%)	72(5.91)	48(9.45)	
MACE(%)	138(11.33)	15(2.95)	< 0.001

Continuous variables are presented as mean ± standard deviation or median (Q1, Q3), as appropriate. Categorical variables are presented as number (percentage).

The derivation cohort was randomly divided into a training set (n = 852, 70%) and an internal validation set (n = 366, 30%). Baseline comparisons between these subsets are shown in Table 2; differences were observed for BB, chloride, and ApoA1.

Table 2
Baseline characteristics according to in-hospital MACE status and derivation subset

Variables	No in-hospital MACE (n = 1080)	In-hospital MACE (n = 138)	Training (n = 852)	Internal validation (n = 366)
SGLT2i	240(22.22)	52(37.68)***	194(22.77)	98(26.78)
CLOP	177(16.39)	32(23.19)*	155(18.19)	54(14.75)
CADH	21(1.94)	1(0.72)	12(1.41)	10(2.73)
LAD	894(82.78)	124(89.86)*	708(83.10)	310(84.70)
HTN	553(51.20)	71(51.45)	419(49.18)	205(56.01)
ATV	312(28.89)	38(27.54)	246(28.87)	104(28.42)
FIB	18(1.67)	1(0.72)	15(1.76)	4(1.09)
SH	477(44.17)	56(40.58)	362(42.49)	171(46.72)
LM	38(3.52)	14(10.14)***	37(4.34)	15(4.10)
ARB	71(6.57)	4(2.90)	55(6.46)	20(5.46)
ASA	1071(99.17)	131(94.93)***	838(98.36)	364(99.45)
LCX	593(54.91)	82(59.42)	463(54.34)	212(57.92)
CAD	121(11.20)	22(15.94)	110(12.91)	33(9.02)
COPD	32(2.96)	5(3.62)	23(2.70)	14(3.83)
RSV	771(71.39)	92(66.67)	607(71.24)	256(69.95)
T2DM	256(23.70)	40(28.99)	207(24.30)	89(24.32)
Stroke	65(6.02)	19(13.77)***	53(6.22)	31(8.47)
Killip	937(86.76)	58(42.03)***	700(82.16)	295(80.60)
RCA	783(72.50)	108(78.26)	624(73.24)	267(72.95)
CCB	102(9.44)	8(5.80)	76(8.92)	34(9.29)
AF	54(5.00)	18(13.04)***	49(5.75)	23(6.28)
DRT	524(48.52)	101(73.19)***	442(51.88)	183(50.00)
ACEI	321(29.72)	18(13.04)***	235(27.58)	104(28.42)
BB	899(83.24)	110(79.71)	709(83.22)	300(81.97)#
Sex	896(82.96)	113(81.88)	704(82.63)	305(83.33)
ARNI	315(29.17)	61(44.20)***	249(29.23)	127(34.70)
TICA	944(87.41)	111(80.43)*	727(85.33)	328(89.62)
DH	286(26.48)	42(30.43)	217(25.47)	111(30.33)

Variables	No in-hospital MACE (n = 1080)	In-hospital MACE (n = 138)	Training (n = 852)	Internal validation (n = 366)
ALT	42.33(19.00, 48.00)	195.89(22.25, 95.25)***	54.48 ± 174.77	71.95 ± 365.68
ALP	83.38(67.00, 96.70)	86.39(66.72, 96.75)	84.42 ± 26.60	82.10 ± 26.47
CKMB	104.82 ± 133.74	120.76 ± 190.34	103.09 ± 130.04	114.86 ± 164.42
DBP	77.82 ± 23.94	70.43 ± 17.14***	76.24 ± 14.44	78.70 ± 36.50
RDW_CV	13.32(12.80, 13.62)	13.65(12.90, 14.20)**	13.38 ± 0.98	13.31 ± 0.92
IT	18.75 ± 42.37	20.57 ± 33.79	20.16 ± 44.61	16.15 ± 32.95
DDimer	1.95 ± 38.08	1.19 ± 2.11	2.42 ± 42.87	0.57 ± 1.18
CHE	8.17(7.06, 9.24)	7.23(5.83, 8.35)***	8.02 ± 1.66	8.18 ± 1.80
Crea	77.07(61.00, 80.00)	114.98(71.00, 121.30)***	79.52 ± 50.41	85.66 ± 80.46
PLCR	35.31 ± 10.96	35.12 ± 9.87	35.38 ± 10.63	35.07 ± 11.32
ECMO	0.32(0.00, 0.00)	17.13(0.00, 0.00)***	2.05 ± 15.15	2.64 ± 17.74
ALB	42.13(40.00, 44.50)	38.32(35.40, 42.38)***	41.55 ± 4.07	42.04 ± 4.17
IG	0.03(0.01, 0.03)	0.12(0.02, 0.09)***	0.04 ± 0.14	0.03 ± 0.05
Ca	2.17 ± 0.57	2.05 ± 0.20*	2.16 ± 0.64	2.15 ± 0.15
LDH	392.57(204, 469.75)	820.77(245.25, 749.25)***	425.13 ± 426.28	478.21 ± 100.13
Glu	7.38(5.09, 8.48)	9.80(6.00, 12.37)***	7.68 ± 3.74	7.58 ± 4.24
GLB	25.92(23.10, 28.40)	24.74(20.73, 27.80)**	25.76 ± 4.34	25.84 ± 5.70
LY_AC	1.48 ± 0.66	1.33 ± 0.95*	1.46 ± 0.71	1.46 ± 0.68
AST_ALT	3.16 ± 5.73	2.77 ± 2.24	2.97 ± 2.16	3.45 ± 9.37
cTnI	7.33 ± 11.74	9.18 ± 12.71	7.40 ± 11.59	7.89 ± 12.47
HR	78.08(67.00, 87.00)	87.47(74.00, 98.00)***	78.96 ± 15.83	79.56 ± 16.19
Lc	1.80(1.06, 2.10)	3.61(1.50, 3.99)***	1.96 ± 2.01	2.11 ± 2.10
MYO	238.91 ± 1003.40	319.39 ± 398.28	253.70 ± 1123.30	234.83 ± 309.65
BMI	24.82 ± 3.52	23.93 ± 3.05**	24.75 ± 3.50	24.65 ± 3.43
MCH	31.80(30.70, 32.90)	31.62(30.02, 32.90)	31.66 ± 2.16	32.06 ± 2.13
NEUT_AC	8.02(5.65, 9.96)	11.50(7.12, 14.52)***	8.41 ± 3.96	8.41 ± 3.73

Variables	No in-hospital MACE (n = 1080)	In-hospital MACE (n = 138)	Training (n = 852)	Internal validation (n = 366)
EO_P	0.67(0.10, 0.80)	0.35(0.00, 0.40)***	0.62 ± 1.24	0.64 ± 1.10
MV_A	0.77 ± 0.16	0.71 ± 0.18***	0.78 (0.66, 0.88)	0.77 (0.65, 0.86)
CK	1120.67(175.75, 1448)	1476.48(190.75, 1809.75)	1162.91 ± 1647.21	1156.51 ± 1525.17
AMY	58.90(43.00, 69.51)	86.12(48.00, 83.75)**	61.50 ± 45.19	63.10 ± 50.41
CO2	21.20(19.38, 23.10)	19.07(17.07, 21.73)***	20.98 ± 3.41	20.91 ± 3.23
HDL-C	1.03(0.88, 1.16)	0.97(0.78, 1.13)**	1.02 ± 0.24	1.04 ± 0.24
MCV	94.01 ± 5.77	94.05 ± 6.45	93.72 ± 5.92	94.70 ± 5.64
TG	1.83 ± 1.56	1.65 ± 1.40	1.79 ± 1.52	1.84 ± 1.59
IP	0.95(0.79, 1.10)	1.26(0.83, 1.43)***	0.98 ± 0.34	0.98 ± 0.42
NEUT_P	77.90(71.75, 85.50)	83.29(78.17, 89.67)***	78.43 ± 9.91	78.70 ± 9.72
IBIL	13.64 ± 6.67	13.03 ± 6.78	13.73 ± 6.82	13.19 ± 6.34
LVEF	53.30(49, 58)	45.31(41.00, 53.00)***	52.42 ± 7.48	52.34 ± 7.43
MCHC	338.32(333, 344)	336.02(330.00, 342.00)*	337.84 ± 7.93	338.58 ± 8.81
AG	15.16(13.10, 17)	18.11(14.40, 19.98)***	15.33 ± 3.69	15.87 ± 4.44
IABP_Duration	0.87(0.00, 0.00)	23.79(0.00, 0.00)***	3.41 ± 21.51	3.59 ± 28.51
RDW_SD	46.60 ± 4.16	47.94 ± 4.98**	46.66 ± 4.32	46.97 ± 4.20
FIB	0.02 ± 0.13	0.01 ± 0.09	0.02 ± 0.13	0.01 ± 0.12
UCR	83.97 ± 26.78	83.61 ± 33.63	84.81 ± 28.75	81.89 ± 24.73
AST	123.64(32, 156.25)	379.98(39.25, 253.75)**	142.20 ± 306.22	177.09 ± 710.51
IG_P	0.25(0.10, 0.30)	0.65(0.18, 0.67)***	0.30 ± 0.61	0.27 ± 0.31
HCY	24.28 ± 18.09	25.82 ± 19.74	24.15 ± 17.52	25.19 ± 19.95
PDW	16.34 ± 0.33	16.40 ± 0.36*	16.34 ± 0.33	16.34 ± 0.34
Age	60.01 ± 12.20	63.25 ± 11.36**	60.78 ± 12.39	59.46 ± 11.53
TC	4.32 ± 1.08	3.89 ± 1.16***	4.25 ± 1.12	4.31 ± 1.05
AGR	1.66(1.50, 1.80)	1.59(1.32, 1.80)**	1.65 ± 0.28	1.66 ± 0.30

Variables	No in-hospital MACE (n = 1080)	In-hospital MACE (n = 138)	Training (n = 852)	Internal validation (n = 366)
PLT	193.18 ± 61.24	199.49 ± 77.62	195.42 ± 65.10	190.34 ± 58.85
CL	105.26(103.2, 107.53)	104.39(102.00, 107.17)*	105.31 ± 3.43	104.83 ± 3.84#
StentNumber	1.33(1.00, 2.00)	1.75(1.00, 2.75)*	1.39 ± 1.09	1.35 ± 1.14
LAD_Echo	3.38(3.10, 3.60)	3.43(3.10, 3.80)	3.39 ± 0.41	3.39 ± 0.44
K	3.77(3.52, 3.99)	4.02(3.58, 4.29)***	3.79 ± 0.45	3.80 ± 0.45
Urea	6.12(4.58, 6.95)	8.50(5.69, 9.58)***	6.40 ± 2.91	6.38 ± 2.95
BA_P	0.24 ± 0.19	0.17 ± 0.17***	0.24 ± 0.20	0.22 ± 0.16
TBA	3.90(1.50, 4.80)	5.69(1.70, 6.00)*	4.10 ± 4.36	4.11 ± 4.83
NTproBNP	1345.95(79, 1091)	4669.03(196.07, 4421.5)***	1524.77 ± 3744.77	2182.63 ± 5609.55
CRP	12.07(1.16, 8.39)	29.75(1.77, 31.57)***	15.89 ± 35.23	9.83 ± 24.17
Height	1.69 ± 0.07	1.68 ± 0.07	1.69 ± 0.07	1.69 ± 0.07
Lpa	24.33 ± 26.72	25.62 ± 31.76	24.30 ± 27.45	24.90 ± 27.05
apoAI	1.18(1.04, 1.32)	1.08(0.89, 1.24)***	1.16 ± 0.23	1.19 ± 0.23#
TBIL	16.94 ± 8.31	17.62 ± 9.21	17.25 ± 8.81	16.48 ± 7.40
SBP	130.42(116, 144)	117.73(98.75, 136.75)***	129.01 ± 22.67	128.92 ± 24.05
PCT	0.21 ± 0.06	0.22 ± 0.07	0.22 ± 0.06	0.21 ± 0.06
GGT	42.40(19.20, 45.60)	61.33(21.25, 71.10)*	43.83 ± 51.31	46.19 ± 69.19
DBIL	3.31(2.10, 4.00)	4.55(2.33, 5.07)***	3.54 ± 2.81	3.23 ± 1.53
UA	353.44(286, 408.25)	398.46(306.73, 475.25)***	355.93 ± 103.54	364.61 ± 111.61
OSM	296.61(293, 300)	301.63(295, 307)***	297.29 ± 6.93	296.92 ± 7.70
AFU	27.96 ± 7.42	27.06 ± 8.52	27.66 ± 7.27	28.32 ± 8.17
Mass	70.79 ± 11.71	67.67 ± 10.81**	70.55 ± 11.72	70.16 ± 11.49
ApoBAI	0.79(0.62, 0.92)	0.82(0.58, 0.99)	0.79 ± 0.29	0.79 ± 0.26
WBC	10.07(7.51, 11.96)	13.52(8.93, 16.74)***	10.47 ± 4.12	10.45 ± 3.92
MPV	11.38 ± 1.58	11.33 ± 1.42	11.38 ± 1.53	11.36 ± 1.63
Hb	150.85(139, 164)	139.05(119.25, 157)***	149.01 ± 22.14	150.69 ± 23.31

Variables	No in-hospital MACE (n = 1080)	In-hospital MACE (n = 138)	Training (n = 852)	Internal validation (n = 366)
LVESV	57.60(46, 65.25)	70.37(53.12, 78.52)***	58.82 ± 19.37	59.58 ± 20.54
EO_AC	0.06 ± 0.11	0.03 ± 0.05*	0.05 ± 0.09	0.06 ± 0.13
LY_P	16.13(9.70, 21.10)	11.24(6.00, 15.38)***	15.60 ± 8.40	15.52 ± 8.28
RBC	4.75(4.35, 5.19)	4.41(3.84, 5.01)***	4.71 ± 0.69	4.71 ± 0.72
αHBDH	349.58(158, 423.75)	536.33(187.25, 639)***	367.72 ± 327.42	377.77 ± 441.46
MV_E	0.68(0.56, 0.79)	0.72(0.56, 0.87)	0.69 ± 0.20	0.68 ± 0.18
HCT	44.59(41.20, 48.33)	41.44(35.98, 46.85)***	44.14 ± 6.46	44.47 ± 6.81
MO_P	5.06 ± 2.03	4.94 ± 2.12	5.10 ± 2.11	4.93 ± 1.89
TP	68.01(64.47, 71.70)	63.09(58.12, 69.17)***	67.32 ± 6.36	67.78 ± 6.77
LDLC	2.86 ± 0.82	2.55 ± 0.86***	2.81 ± 0.85	2.86 ± 0.78
MO_AC	0.49(0.33, 0.60)	0.65(0.40, 0.81)***	0.52 ± 0.29	0.50 ± 0.23
BA_AC	0.02 ± 0.02	0.02 ± 0.02**	0.02 ± 0.02	0.02 ± 0.02
Mg	0.83(0.78, 0.88)	0.88(0.78, 0.93)**	0.83 (0.78, 0.89)	0.83 (0.78, 0.89)
LVEDV	122.86 ± 31.48	126.89 ± 34.00	122.40 ± 26.89	125.46 ± 40.96
apoB	0.91 ± 0.27	0.83 ± 0.29**	0.89 ± 0.28	0.91 ± 0.26

Compared with the no-MACE group, *P < 0.05; **P < 0.01; ***P < 0.001.

Compared with the training set, #P < 0.05; ##P < 0.01; ###P < 0.001.

Continuous variables are presented as mean ± standard deviation or median (Q1, Q3), as appropriate. Categorical variables are presented as number (percentage).

Model Development

We first compared alternative preprocessing strategies in the internal validation cohort (Fig. 2 and Table 3). Boruta+GSCV achieved an AUROC of 0.83 (95% CI 0.74–0.90) and an AUPRC of 0.67 (95% CI 0.52–0.79), while LASSO+GSCV achieved an AUROC of 0.82 (95% CI 0.73–0.90) and an AUPRC of 0.64 (95% CI 0.49–0.77). ADASYN-containing pipelines produced similar AUROC values but lower AUPRC values (0.557–0.561) than Boruta+GSCV (0.674), indicating that synthetic oversampling did not provide a clear overall advantage in this dataset.

Table 3
Performance of fusion model based on different data preprocessing strategies

Data Processing	Accuracy	Precision	Sensitivity	Specificity	F1-score	AUROC	AUPRC
Boruta+GSCV	0.932	0.846	0.512	0.988	0.638	0.827	0.674
LASSO+GSCV	0.923	0.778	0.488	0.981	0.600	0.821	0.638
ADASYN+GSCV	0.923	0.857	0.419	0.991	0.563	0.827	0.561
ADASYN+LASSO+GSCV	0.910	0.708	0.395	0.978	0.507	0.829	0.557
ADASYN+Boruta+GSCV	0.913	0.789	0.349	0.988	0.484	0.834	0.560

Abbreviations: ADASYN, adaptive synthetic sampling; LASSO: least absolute shrinkage and selection operator; GSCV: grid search cross-validation; RF: random forest; AUROC, area under the receiver operating characteristic curve; AUPRC, area under the precision-recall.

Table 4
Model performance based on ML algorithms

Algorithms	Accuracy	Precision	Sensitivity	Specificity	F1-score	AUROC	AUPRC
RF	0.932	0.875	0.488	0.991	0.627	0.809	0.653
LightGBM	0.891	0.579	0.256	0.975	0.355	0.680	0.378
SVM	0.885	1.000	0.023	1.000	0.045	0.758	0.591
XGBoost	0.902	0.889	0.186	0.997	0.308	0.784	0.650
LR	0.902	0.581	0.581	0.944	0.581	0.754	0.604
KNN	0.932	0.875	0.488	0.991	0.627	0.776	0.627
DNN	0.907	0.846	0.256	0.994	0.393	0.767	0.588

Abbreviations: ML, machine learning; GSCV, grid search cross-validation; RF, random forest; LR, logistic regression; LightGBM, light gradient boosting machine; SVM, support vector machine; XGBoost, extreme gradient boosting; KNN, k-nearest neighbors; DNN, deep neural network; AUROC, area under the receiver operating characteristic curve; AUPRC, area under the precision-recall.

Bootstrap comparisons did not show a statistically significant difference between the leading Boruta- and LASSO-based preprocessing strategies. Sequential LASSO and Boruta selection was therefore used to obtain a compact predictor set, while RF was selected as the final learner based on the balance of discrimination, AUPRC, parsimony, and interpretability. ADASYN was evaluated but was not retained as an essential component of the final pipeline because it did not improve AUPRC.

Feature selection proceeded in two stages. LASSO reduced the 125 candidate variables to nine, and Boruta-based refinement retained eight variables (Figure 4): Killip class, ECMO, inorganic phosphorus (IP), absolute neutrophil count (NEUT_AC), total protein (TP), LVEF, ALT, and potassium (K).

When comparing different modeling algorithms, we found that they varied in their ability to select predictive features. Using the variables selected by LASSO and Boruta, we built prediction models with various ML

algorithms. The results showed that LR and DNN mainly adjusted feature weights without changing the number of predictors, whereas the remaining algorithms both reduced the number of features and reassigned their weights (Figure 5).

The feature radar chart (Figure 6A) and bar chart (Figure 6B) present the feature names and their weight rankings in the RF model for predicting In-hospital MACE in post-PCI ACS patients. The eight included features are ranked by weight as follows: Killip (0.377), IP (0.156), ECMO (0.128), LVEF (0.087), NEUT_AC (0.077), TP (0.077), ALT (0.053), and K (0.045). For localized interpretability of individual predictions, we applied the LIME method (Figure 6C), which illustrates, for a specific patient, each predictor's contribution to the risk assessment. In this case, ECMO was the most influential predictor, followed by IP and Killip. Specifically, protective factors included absence of ECMO, 48% <LVEF< 53.5%, Killip class I, 32< ALT≤49.25, 67.7< TP≤ 71.4, and 3.52<K≤3.76, whereas IP >1.13 and NEUT_AC >10.36 were identified as risk-aggravating factors. To further elucidate the contribution of individual features to the overall model, SHAP analysis was conducted. The resulting feature importance scatter plot (Figure 6D) delineated the ranking as: Killip, IP, ECMO, LVEF, TP, NEUT_AC, K, and ALT.

The calibration curve showed reasonable agreement between predicted and observed risks across much of the evaluated range (Figure 7A). Decision-curve analysis suggested greater net benefit than treat-all and treat-none strategies across a range of threshold probabilities (Figure 7B). The RF model achieved an AUROC of 0.82 (95% CI 0.73-0.90) in the internal validation cohort (Figure 7C) and 0.90 (95% CI 0.85-0.95) in the external validation cohort (Figure 7D). Because only 15 MACE events occurred in the external cohort and its event rate differed substantially from that of the derivation cohort, the external AUROC should be interpreted primarily as evidence of preserved ranking ability rather than definitive proof of calibration or broad transportability.

Using Streamlit, we implemented a prototype interface for the final RF model (Figure 8). The application accepts the eight model variables and returns an estimated probability of in-hospital MACE together with a local explanation. For the illustrative case shown, the estimated risk was 25.7%. This prototype demonstrates technical deployability but was not prospectively evaluated for clinical impact or workflow integration.

Discussion

This two-center study developed and externally evaluated a parsimonious machine-learning model for in-hospital MACE after PCI in ACS. Starting from 125 routinely collected EMR variables, sequential LASSO and Boruta selection yielded an eight-variable RF model. The model showed good discrimination in internal validation (AUROC 0.82) and preserved discrimination in an independent hospital cohort (AUROC 0.90). The principal contribution is therefore not the use of RF alone, but the combination of broad EMR screening, explicit comparison of preprocessing strategies, aggressive reduction to a small predictor set, model interpretation, and independent-center validation.

The preprocessing experiments provide an important methodological observation. ADASYN-containing pipelines produced AUROC values similar to non-ADASYN strategies but lower AUPRC than Boruta+GSCV. This is relevant because AUROC can appear reassuring in imbalanced datasets even when positive-class performance remains limited[15, 17, 20, 21, 22, 23]. Our findings therefore support evaluating oversampling empirically rather than assuming that class balancing will improve clinical prediction. Reporting both AUROC and AUPRC also provides a more complete description of model behavior for a relatively uncommon outcome.

Model parsimony was a second design priority. Sequential LASSO and Boruta screening reduced 125 candidate variables to eight without an obvious loss of discrimination. Compact models may be easier to audit, transport, and integrate into clinical information systems than models requiring dozens of inputs. Nevertheless, parsimony does not by itself establish clinical usefulness; calibration, threshold-specific performance, workflow timing, and prospective impact remain necessary considerations. The literature also illustrates the diversity of populations and outcomes used for cardiovascular prediction: population-based cardiovascular incidence [24, 25], post-PCI readmission in non-ST-elevation myocardial infarction [26], external validation of ACS risk scores after PCI [27], serum biomarker prognosis in ACS [28], cardiorenal risk prediction in ACS [29], and angiographic no-reflow after primary PCI [30, 31]. These studies address different endpoints or clinical settings and are not direct benchmarks for the present in-hospital composite MACE model.

The final RF model had high specificity (0.991) but only moderate sensitivity (0.488) at the reported operating point. Accordingly, it should not be used as a stand-alone screening tool to exclude risk. Thresholds would need to be selected according to the intended clinical action and the relative consequences of false-negative and false-positive classifications. Decision-curve analysis suggested potential net benefit across a range of thresholds, but this finding requires confirmation in prospective implementation studies.

The eight retained predictors also have plausible clinical interpretations. Killip class and LVEF reflect hemodynamic and ventricular functional status, while NEUT_AC may capture acute inflammatory stress. Electrolyte and biochemical variables, including IP, K, TP, and ALT, may reflect metabolic, nutritional, hepatic, or systemic consequences of severe ACS. ECMO use is strongly associated with critical illness and therefore functions primarily as a marker of extreme clinical severity rather than a causal risk factor. The SHAP and LIME analyses improve transparency by showing how these variables contribute to model output, but they do not establish causal relationships.

Limitations

This study has several limitations. First, it was retrospective and involved only two hospitals, which may introduce selection bias and limit generalizability. Second, the external cohort contained only 15 MACE events and had a substantially lower event rate than the derivation cohort; external calibration and subgroup performance therefore require confirmation in larger datasets. Third, the temporal availability of all predictors was not standardized to a single prespecified prediction time; in particular, LVEF was obtained within 24 hours and ECMO may reflect evolving clinical severity. Fourth, the composite MACE outcome combines clinically heterogeneous events, and event-specific models were not feasible with the available sample size. Fifth, the model was not directly compared with established clinical scores such as GRACE or PRAISE. Finally, the Streamlit application is a prototype and has not undergone prospective impact, usability, or workflow evaluation.

Conclusion

A parsimonious eight-variable RF model derived from routinely collected EMR data showed good internal discrimination for in-hospital MACE after PCI in ACS and preserved discrimination in an independent hospital cohort. Sequential feature selection reduced model complexity, while comparison of preprocessing strategies showed no clear benefit from synthetic oversampling. These findings support further evaluation of the model as an early in-hospital clinical decision-support approach. Prospective multicenter validation with a prespecified

prediction time, direct comparison with established risk scores, and formal calibration assessment is required before clinical implementation.

Abbreviations

ACEF: the age-creatinine-ejection fraction score

Adaboost: adaptive boosting

ADASYN: adaptive synthetic sampling

AMI: acute myocardial infarction

AUROC: area under the receiver operating characteristic curve

AUPRC: area under the precision-recall curve

CAD: coronary artery disease

DNN: deep neural network

GSCV: grid search with cross-validation

KNN: k-nearest neighbors

LASSO: least absolute shrinkage and selection operator

LightGBM: light gradient boosting machine

LIME: local interpretable model-agnostic explanations

LR: logistic regression

MICE: multiple imputation by chained equations

ML: machine learning

SHAP: shapley additive explanation

PCI: percutaneous coronary intervention

XGBoost: extreme gradient boosting

Declarations

Ethics approval and consent to participate

All procedures were performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. This study was approved by the Ethics Committee of the First Hospital of

Lanzhou University (Approval No. LDYYLL-2026-801) and Gansu Gem Flower Hospital (Approval No. LZSH-IEC-2026-109-1), which waived the requirement for individual consent due to the use of retrospective data.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the following grants: the Natural Science Foundation of Gansu Province (No. 26JRRA301); "Multi-modal Adverse Events Prediction for Premature Coronary Heart Disease Trial (MAP-CHD)" (No. 2023ZD0504003); the National Key R&D Program of China (No. 2018YFC1311505); and the Lanzhou Science and Technology Bureau (No. 2023-2-119).

Authors' contributions

DY and ZZ designed the research study. WT, LT collected the patient data. EX and YW performed the research. WL analyzed the data. DY wrote the manuscript. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Kraler S, Mueller C, Libby P et al. Acute coronary syndromes: mechanisms, challenges, and new opportunities.[J]European Heart Journal. 2025, (46), 2866–89.
2. Lüsebrink EBinzenhöferL, Adamo M et al. Cardiogenic shock.[J]The Lancet. 2024, (404), 2006–20.
3. Chan K, Wahome E, Tsiachristas A et al. Inflammatory risk and cardiovascular events in patients without obstructive coronary artery disease: the ORFAN multicentre, longitudinal cohort study.[J]The Lancet. 2024, (403), 2606–18.
4. Zaka A, Mutahar D, Gorcilov J et al. Machine learning approaches for risk prediction after percutaneous coronary intervention: a systematic review and meta-analysis.[J]European Heart. J - Digit Health. 2025, (6), 23–44.
5. Fan Z, Du Z, Fu J et al. Comparing the performance of machine learning and conventional models for predicting atherosclerotic cardiovascular disease in a general Chinese population.[J]BMC Medical

Informatics and Decision Making. 2023, (23).

6. Han B, Zhang Z, He S et al. Machine Learning Approach on Predictive Model Establishment for In-Hospital Mortality in Acute Myocardial Infarction Patients Post-Percutaneous Coronary Intervention: Solutions for Databases With Dimensionality Reduction and Class Imbalance.[J]Reviews in Cardiovascular Medicine. 2025, (26).
7. Du S, Li W, Wang Y et al. Development and validation of an in-hospital cardiogenic shock prediction model for AMI patients based on machine learning.[J]BMC Cardiovascular Disorders. 2026, (26).
8. Simon S, Mandair D, Albakri A et al. The Impact of Time Horizon on Classification Accuracy: Application of Machine Learning to Prediction of Incident Coronary Heart Disease.[J]JMIR Cardio. 2022, (6).
9. Zworth M, Kareemi H, Boroumand S et al. Machine learning for the diagnosis of acute coronary syndrome using a 12-lead ECG: a systematic review.[J]Canadian. J Emerg Med. 2023, (25), 818–27.
10. Emakhu J, Monplaisir L, Aguwa C et al. Acute coronary syndrome prediction in emergency care: A machine learning approach.[J]Computer Methods and Programs in Biomedicine. 2022, (225).
11. Hu J, Xu J, Li M et al. Identification and validation of an explainable prediction model of acute kidney injury with prognostic implications in critically ill children: a prospective multicenter cohort study. [J]eClinicalMedicine. 2024, (68).
12. D'Ascenzo F, De Filippo O, Gallone G et al. Machine learning-based prediction of adverse events following an acute coronary syndrome (PRAISE): a modelling study of pooled datasets.[J]The Lancet. 2021, (397), 199–207.
13. Cumitini L, Giubertoni A, Rossi L et al. Stratification of Early Arrhythmic Risk in Patients Admitted for Acute Coronary Syndrome: The Role of the Machine Learning-Derived PRAISE Score.[J]Clinical Cardiology. 2024, (47).
14. Georgiopoulos G, Kraler S, Mueller-Hennessen M et al. Modification of the GRACE Risk Score for Risk Prediction in Patients With Acute Coronary Syndromes.[J]JAMA Cardiology. 2023, (8).
15. Greco A, Capodanno D. Personalized Treatment of Patients with Coronary Artery Disease: The Value and Limitations of Predictive Models.[J]Journal of Cardiovascular Development and Disease. 2025, (12).
16. Zhang X, Wang X, Xu L et al. The predictive value of machine learning for mortality risk in patients with acute coronary syndromes: a systematic review and meta-analysis.[J]European Journal of Medical Research. 2023, (28).
17. Sousa S, Paredes S, Rocha T et al. Machine learning models' assessment: trust and performance.[J]Medical & Biological Engineering & Computing. 2024, (62), 3397–3410.
18. Khera R, Oikonomou EK, Nadkarni GN et al. Transforming Cardiovascular Care With Artificial Intelligence: From Discovery to Practice.[J]Journal of the American College of Cardiology. 2024, (84), 97–114.
19. Wenzl FA, Ow K W,Velders MA et al. Prediction of mortality, bleeding, and ischaemic events in patients with cancer and acute coronary syndrome: a model development and validation study.[J]The Lancet. 2026, (407), 515–28.
20. Chen C, Zhang J, Zhang T et al. Application of Machine Learning Algorithms in Predicting Major Adverse Cardiovascular Events after Percutaneous Coronary Intervention in Patients with New-Onset ST-Segment Elevation Myocardial Infarction.[J]Reviews in Cardiovascular Medicine. 2025, (26).

21. Bouzid Z, Sejdic E, Martin-Gill C et al. Electrocardiogram-based machine learning for risk stratification of patients with suspected acute coronary syndrome[. J]European Heart J. 2025, (46), 943–54.
22. Zheng B, Liang T, Mei J et al. Prediction of 90 Day Readmission in Heart Failure with Preserved Ejection Fraction by Interpretable Machine Learning.[J]ESC Heart Failure. 2024, (11), 4267–76.
23. Li WYD, Hu W et al. Enhancing one-year mortality prediction in STEMI patients post-PCI: an interpretable machine learning model with risk stratification.[J]Frontiers in Artificial Intelligence. 2025, (8).
24. Gao J, Li C-p, Liu Y et al. Using Machine Learning to Predict MACEs Risk in Patients with Premature Myocardial Infarction.[J]Reviews in Cardiovascular Medicine. 2025, (26).
25. Huang Q, Jiang Z, Shi B et al. Characterisation of cardiovascular disease (CVD) incidence and machine learning risk prediction in middle-aged and elderly populations: data from the China health and retirement longitudinal study (CHARLS).[J]BMC Public Health. 2025, (25).
26. Liu Y, Du L, Li L et al. Development and validation of a machine learning-based readmission risk prediction model for non-ST elevation myocardial infarction patients after percutaneous coronary intervention. [J]Scientific Reports. 2024, (14).
27. Molenaar MA, Selder JL, Schmidt AF et al. Validation of machine learning-based risk stratification scores for patients with acute coronary syndrome treated with percutaneous coronary intervention.[J]European Heart. J - Digit Health. 2024, (5), 702–11.
28. Zuo G, Zhang J, Xie H. Prognostic value of serum angiopoietin-like protein 2 in patients with acute coronary syndrome.[J]BMC Cardiovascular Disorders. 2024, (24).
29. Wenzl FA, Wang P, Arrigo M et al. Proenkephalin improves cardio-renal risk prediction in acute coronary syndromes: the KID-ACS score.[J]European Heart Journal. 2025, (46), 38–54.
30. Zhao Y, Yang J, Ji Y et al. Usefulness of fibrinogen-to-albumin ratio to predict no-reflow and short-term prognosis in patients with ST-segment elevation myocardial infarction undergoing primary percutaneous coronary intervention.[J]Heart Vessels. 2019, (34), 1600–7.
31. Zhang H, Qiu B, Zhang Y et al. The Value of Pre-Infarction Angina and Plasma D-Dimer in Predicting No-Reflow After Primary Percutaneous Coronary Intervention in ST-Segment Elevation Acute Myocardial Infarction Patients.[J]Medical Science Monitor. 2018, (24), 4528–35.