

Development and Validation of a Nomogram for Predicting Cardiogenic Shock in Elderly Patients with ST-Segment Elevation Myocardial Infarction: A Retrospective Cohort Study

Zunguo Liu

General Hospital of Ningxia Medical University

Shaobing Yang

General Hospital of Ningxia Medical University

Yuanyuan Mao

General Hospital of Ningxia Medical University

Li Ma

General Hospital of Ningxia Medical University

Wanyu Song

Ningxia Medical University

Hui Huang

huanghui862@163.com

General Hospital of Ningxia Medical University

Research Article

Keywords: ST-elevation myocardial infarction, cardiogenic shock, elderly, nomogram, neutrophil-to-lymphocyte ratio, D-dimer

Posted Date: October 18th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10895545/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

Cardiogenic shock (CS) remains the most lethal complication of ST-elevation myocardial infarction (STEMI), particularly in elderly patients who lack dedicated prediction tools. This study developed and internally validated a nomogram for admission-time CS risk stratification in elderly (≥ 60 years) STEMI patients, and evaluated the incremental value of combining Killip classification with laboratory parameters.

Methods

This single-center retrospective cohort study consecutively enrolled 527 patients aged ≥ 60 years with STEMI between January 2019 and December 2022; after excluding 4 patients with missing CS status, 523 were analyzed (CS, $n = 39$; non-CS, $n = 484$). Admission demographics, Killip classification, comorbidities, vital signs, laboratory tests, echocardiography, and coronary angiography data were collected. Candidate variables were screened by univariate logistic regression ($P < 0.1$), followed by multivariate logistic regression. Discrimination was assessed using the area under the receiver operating characteristic curve (AUC), calibration by the Hosmer-Lemeshow test, and internal validation by 1000 bootstrap resamples; the likelihood ratio test and DeLong test compared models, and decision curve analysis evaluated clinical utility.

Results

Killip class III-IV (OR = 13.531, 95% CI: 5.541–33.039, $P < 0.001$) and age (OR = 1.076, 95% CI: 1.016–1.139, $P = 0.013$) were independent predictors of CS. The combined model (Killip class, D-dimer, neutrophil-to-lymphocyte ratio, left ventricular ejection fraction, glucose, and age) achieved an AUC of 0.874 (95% CI: 0.819–0.927), significantly outperforming Killip classification alone (AUC = 0.815, DeLong $P = 0.001$) and the laboratory-only model (AUC = 0.775, DeLong $P = 0.014$). A parsimonious Killip + age model achieved a comparable AUC of 0.864 (likelihood ratio test $P = 0.168$; DeLong $P = 0.140$). The bootstrap-corrected AUC was 0.858 with good calibration (Hosmer-Lemeshow $P = 0.412$). Decision curve analysis demonstrated positive net benefit across the 5%-25% threshold range, and a simplified 25-point bedside scoring rule achieved an AUC of 0.846.

Conclusions

The nomogram effectively predicts admission-time CS risk in elderly STEMI patients. When Killip classification is available, adding age alone yields a powerful and parsimonious bedside tool, while the

full nomogram remains preferable when maximal predictive performance is pursued. External validation is required.

Background

ST-elevation myocardial infarction (STEMI) remains one of the most critical cardiovascular emergencies. Although the widespread adoption of primary percutaneous coronary intervention (PCI) has significantly improved overall outcomes, cardiogenic shock (CS) remains the most common and lethal complication of STEMI. The incidence of CS in STEMI patients ranges from 6% to 10%, with in-hospital mortality rates as high as 40%–50%.^[1, 2] A 2024 Lancet review highlighted that although the proportion of AMI-related CS has declined, AMI remains a major cause of CS, accounting for 30%–81% of all CS cases.^[3] Early identification of patients at high risk for CS helps guide timely initiation of mechanical circulatory support (MCS), vasoactive medications, and optimization of revascularization strategies.

Elderly STEMI patients represent a particularly vulnerable population. With advancing age, patients often present with multiple comorbidities, more complex coronary anatomy, and atypical clinical manifestations, leading to higher CS incidence and worse outcomes.^[4] Previous studies have reported CS rates of 15%–20% in elderly AMI patients aged ≥ 75 years, significantly higher than in younger patients.^[5] The 2022 Society for Cardiovascular Angiography and Interventions (SCAI) updated CS classification system (Stages A–E) emphasizes that intervention before CS progresses to Stage C (classic CS) is critical for improving outcomes.^[6] However, elderly patients are particularly prone to being overlooked during the early, subclinical phases of CS. Dedicated prediction tools for CS risk in elderly STEMI patients remain limited.

The Killip classification is a classic bedside tool for assessing cardiac function in AMI, with Killip classes III–IV closely associated with CS. However, Killip classification primarily reflects the clinical status at presentation, and its sensitivity and specificity for CS prediction are limited when used alone. In recent years, several hematological markers have been associated with adverse outcomes in AMI. The neutrophil-to-lymphocyte ratio (NLR), a systemic inflammatory marker, was shown in a 2024 meta-analysis of 37 studies to be associated with a 2.29-fold increase in all-cause mortality among MI patients.^[7] D-dimer, reflecting coagulation-fibrinolysis activation, has been independently associated with in-hospital adverse events in AMI patients.^[8] Left ventricular ejection fraction (LVEF), hemoglobin, and glucose levels have also been linked to CS development.^[9, 10] Combining these laboratory parameters with clinical assessment may improve CS risk prediction.

A nomogram is a graphical tool that visualizes multivariate regression models, allowing convenient calculation of individual event probabilities, and has been widely applied in prognostic assessment across various diseases.^[11] Recently, Fang et al.^[12] developed a CS survival prediction nomogram (AUC = 0.76) based on the MIMIC-IV database; however, this model targeted patients who had already

developed CS in the ICU, rather than predicting CS risk. Another study^[13] constructed a CS prediction model for AMI patients using LASSO regression but did not specifically focus on the elderly population.

This study aimed to develop and validate a nomogram for overall admission-time CS risk stratification in elderly STEMI patients based on admission clinical parameters, and to evaluate the incremental predictive value of combining Killip classification with laboratory parameters. In this regard, two distinct clinical questions should be distinguished: question A, the assessment of overall CS status and risk at admission (including patients who already present in shock), which the present study can directly answer; and question B, the prediction of *de novo* CS progression in patients without shock at presentation (Killip class I–III), which cannot be adequately answered with the present sample and requires dedicated, adequately powered cohorts.

Methods

Study design and population

This single-center retrospective cohort study consecutively enrolled STEMI patients hospitalized at the Department of Cardiology of the General Hospital of Ningxia Medical University between January 2019 and December 2022. Inclusion criteria were: (1) age ≥ 60 years; (2) confirmed STEMI diagnosis based on typical chest pain, electrocardiographic ST-segment elevation, and elevated cardiac biomarkers.^[14] Exclusion criteria were: (1) missing key data (CS status, Killip classification) at admission; (2) non-ST-elevation myocardial infarction. The study was approved by the Ethics Committee of the General Hospital of Ningxia Medical University, with informed consent waived due to the retrospective design and anonymized data. The study adhered to the principles of the Declaration of Helsinki.

Data collection

Data were extracted from the electronic medical record system:

- - **Demographics:** Age, sex
- - **Clinical assessment:** Killip classification (I–IV), admission heart rate and blood pressure
- - **Comorbidities:** Hypertension, diabetes, chronic obstructive pulmonary disease (COPD), chronic kidney disease (CKD), prior PCI
- - **Laboratory tests:** White blood cell count (WBC), neutrophil percentage, lymphocyte percentage, monocyte percentage, hemoglobin, platelet count, low-density lipoprotein cholesterol, fasting glucose, serum creatinine, estimated glomerular filtration rate (eGFR), troponin, NT-proBNP, potassium, sodium, calcium, fibrinogen, D-dimer, albumin, aspartate aminotransferase, alanine aminotransferase
- - **Echocardiography:** Left ventricular end-diastolic diameter (LVEDD), LVEF
- - **Coronary angiography:** Chronic total occlusion (CTO) lesion, three-vessel disease

- - **Electrocardiography:** QRS duration, corrected QT interval
- - **Calculated indices:** NLR (neutrophil/lymphocyte), systemic immune-inflammation index (SII, platelet $\times$ NLR), platelet-to-lymphocyte ratio (PLR, platelet/lymphocyte)

All laboratory values were from the first measurement within 24 hours of admission.

Definitions

Cardiogenic shock: Diagnosed based on the following criteria:^[15] (1) systolic blood pressure < 90 mmHg for ≥ 30 minutes or requiring vasopressors; (2) signs of tissue hypoperfusion (cool extremities, decreased urine output, altered mental status); (3) exclusion of other causes of hypotension. This definition is consistent with SCAI CS classification Stage C and above.^[6]

- - Killip classification: Class I: no signs of heart failure; Class II: rales $< 50\%$ of lung fields, S3 gallop; Class III: acute pulmonary edema; Class IV: cardiogenic shock.^[16] Killip classes III–IV were combined as a binary variable for analysis.
- - **Elderly:** Age ≥ 60 years (per WHO Western Pacific regional definition).

Statistical analysis

Sample size assessment. A total of 527 patients were enrolled; after excluding 4 patients with missing cardiogenic shock status, 523 patients were analyzed, including 39 CS events. Based on the events-per-variable (EPV) ratio ≥ 6 principle,^[17] 39 events can support approximately 6 predictor variables, meeting the modeling requirement.

Descriptive statistics. Continuous variables were expressed as median (interquartile range, IQR) and compared using the Mann-Whitney U test. Categorical variables were expressed as frequency (percentage) and compared using chi-square or Fisher's exact test (when any expected cell frequency was < 5).

Variable selection and model construction. Univariate logistic regression first screened candidate variables associated with CS ($P < 0.1$). To avoid multicollinearity, highly correlated variables were selectively removed: NLR was retained over individual neutrophil and lymphocyte percentages; systolic blood pressure was retained over diastolic; SII and PLR were removed in favor of NLR. Final model variables were determined considering EPV and clinical importance. LASSO regression (L1-penalized logistic regression) was also applied to all $P < 0.1$ candidate variables as a sensitivity analysis for variable selection. Variance inflation factors (VIF) were calculated to assess multicollinearity (VIF > 5 indicates significant collinearity).

Two multivariate logistic regression models were constructed: (1) Combined model: Killip III–IV, D-dimer, NLR, LVEF, glucose, and age; (2) Laboratory-only model: D-dimer, NLR, LVEF, glucose, age, and

hemoglobin (without Killip classification). Iterative reweighted least squares estimation was used, reporting odds ratios (OR) with 95% confidence intervals (CI). To directly address the parsimony of the combined model, a reduced model restricted to Killip class III–IV and age was also fitted, and the likelihood ratio test (LRT) was used to compare the nested models (combined vs. reduced), with the DeLong test used for non-nested AUC comparisons.

Model evaluation. (1) Discrimination: area under the receiver operating characteristic curve (AUC) with 95% CI; DeLong test compared AUCs across models. (2) Calibration: Hosmer-Lemeshow goodness-of-fit test. (3) Internal validation: 1000 bootstrap resamples to calculate optimism-corrected AUC (corrected AUC = original AUC – mean optimism). (4) Collinearity assessment: VIF. (5) Sensitivity analyses: Firth penalized logistic regression (for rare events); NLR×Killip interaction analysis; stepwise forward regression (Akaike information criterion, AIC) as an alternative to LASSO for variable selection. (6) Clinical utility: decision curve analysis (DCA) for net benefit assessment. (7) Simplified scoring: integer-based clinical scoring rule derived from regression coefficients. ROC analysis with the Youden index determined optimal cutoffs for continuous predictors. Because the optimal cutoffs were derived from the same derivation cohort, they are subject to optimism bias and require external validation.

Handling of extreme NLR values. Thirteen cases with NLR > 20 (including 5 cases with NLR > 30) were identified. Upon verification, these extreme values were physiologically plausible (corresponding to neutrophil percentage > 90% and lymphocyte percentage < 3% in severe stress states). NLR was included as a continuous variable without transformation in the primary analysis, as logistic regression is relatively robust to extreme values in continuous predictors. A sensitivity analysis excluding NLR > 30 cases showed no material change in model performance (AUC change < 0.01), confirming that extreme values did not materially influence the results.

Missing data handling. All variables had missing rates < 10% (highest: troponin at 8.9%). Complete-case analysis (listwise deletion) was used, yielding 513/527 (97.3%) complete cases. The missing pattern was consistent with missing at random, and sensitivity analyses showed no material impact on results. As an additional safeguard, multiple imputation (chained equations, 20 imputations) was considered; however, given the low overall missing rate and the absence of a monotone missing pattern related to outcome, complete-case analysis was retained as the primary approach, with the caveat that non-random missingness of troponin (8.9%) cannot be fully excluded.

Two-sided $P < 0.05$ was considered statistically significant. Analyses were performed using Python 3.13.

Results

Baseline characteristics

A total of 527 elderly STEMI patients were enrolled; 4 were excluded due to missing cardiogenic shock status, leaving 523 patients for analysis, including 39 (7.5%) in the CS group and 484 (92.5%) in the non-

CS group. The overall median age was 70 years (IQR: 65–76), with males comprising 73.6%. Among the 523 patients, 107 (20.5%) were classified as Killip III–IV.

CS patients were older (71 vs. 69 years, $P = 0.006$), had higher heart rates (85.5 vs. 78.0 bpm, $P = 0.011$), and lower systolic blood pressure (109 vs. 120 mmHg, $P = 0.019$). Laboratory findings showed that CS patients had significantly elevated D-dimer (1.51 vs. 0.59 mg/L, $P < 0.001$), NT-proBNP (2490 vs. 807 pg/mL, $P = 0.002$), glucose (6.79 vs. 6.22 mmol/L, $P = 0.002$), and WBC (12.01 vs. 9.61×10^9 /L, $P = 0.005$), while hemoglobin (129 vs. 140.5 g/L, $P < 0.001$), eGFR (87.7 vs. 102.9 mL/min/1.73m², $P = 0.002$), LVEF (45.0% vs. 48.7%, $P < 0.001$), and albumin (35.4 vs. 37.1 g/L, $P = 0.012$) were significantly lower. NLR showed a trend toward elevation in the CS group (6.19 vs. 5.32, $P = 0.054$).

Killip class III–IV was present in 79.5% of CS patients versus 15.7% of non-CS patients ($P < 0.001$). No significant differences were observed in sex, hypertension, diabetes, COPD, CKD, or other comorbidities between groups. Baseline characteristics are presented in Table 1.

Table 1
Baseline characteristics of elderly STEMI patients by cardiogenic shock status.

Variable	Non-CS (*n* = 484)	CS (*n* = 39)	*P*-value
Demographics			
Age, years	69.0 (65.0–75.0)	71.0 (69.0–79.0)	0.006
Male, *n* (%)	358 (74.0)	27 (69.2)	0.501
Clinical assessment			
Killip III–IV, *n* (%)	76 (15.7)	31 (79.5)	< 0.001
Heart rate, bpm	78.0 (68.0–89.0)	85.5 (74.0–102.0)	0.011
Systolic BP, mmHg	120.0 (107.0–134.0)	109.0 (90.0–133.0)	0.019
Diastolic BP, mmHg	74.0 (67.0–86.0)	68.5 (60.0–78.0)	0.028
Comorbidities, *n* (%)			
Hypertension	274 (56.6)	23 (60.5)	0.639
Diabetes	134 (27.7)	8 (21.6)	0.421
COPD	20 (4.1)	2 (5.3)	0.671†
CKD	26 (5.4)	2 (5.3)	1.000†
Prior PCI	36 (7.5)	3 (7.9)	0.757†
Laboratory tests			
WBC, ×10 ⁹ /L	9.61 (7.79–11.89)	12.01 (9.29–17.56)	0.005
Neutrophil, %	77.3 (67.1–84.2)	80.8 (73.4–88.0)	0.016
Lymphocyte, %	14.5 (9.3–21.6)	12.7 (7.0–19.5)	0.089
Hemoglobin, g/L	140.5 (127.0–151.0)	129.0 (111.0–145.0)	< 0.001
Platelet, ×10 ⁹ /L	214.0 (176.0–259.0)	182.0 (149.0–231.0)	0.025
Glucose, mmol/L	6.22 (5.31–7.81)	6.79 (6.12–9.28)	0.002
eGFR, mL/min/1.73m ²	102.9 (82.6–119.2)	87.7 (39.8–110.3)	0.002
Creatinine, μmol/L	67.7 (58.0–82.4)	74.3 (61.8–120.0)	0.006
Troponin, ng/mL	6.91 (1.62–20.0)	17.66 (3.60–25.12)	0.029
NT-proBNP, pg/mL	806.8 (260.6–2673.2)	2490.1 (863.0–5928.0)	0.002
Calcium, mmol/L	2.14 (2.09–2.23)	2.11 (2.03–2.14)	0.006

Variable	Non-CS (*n* = 484)	CS (*n* = 39)	*P*-value
D-dimer, mg/L	0.59 (0.28–1.12)	1.51 (0.56–3.92)	< 0.001
Albumin, g/L	37.1 (34.3–39.8)	35.4 (31.4–38.0)	0.012
Echocardiography & Angiography			
LVEF, %	48.7 (45.1–56.3)	45.0 (42.7–49.4)	< 0.001
LVEDD, mm	52.0 (50.0–56.0)	51.0 (47.0–56.0)	0.269
CTO lesion, *n* (%)	108 (22.6)	11 (31.4)	0.232
Three-vessel disease, *n* (%)	242 (52.4)	14 (45.2)	0.436
Inflammatory indices			
NLR	5.32 (3.03–8.80)	6.19 (3.65–11.62)	0.054

Continuous variables are expressed as median (IQR), compared using Mann-Whitney U test; categorical variables as frequency (%), compared using chi-square test. †Fisher's exact test. Of the 527 enrolled patients, 4 were excluded due to missing cardiogenic shock status; percentages are based on the 523 patients in the final analysis cohort. BP: blood pressure; CKD: chronic kidney disease; COPD: chronic obstructive pulmonary disease; CS: cardiogenic shock; CTO: chronic total occlusion; eGFR: estimated glomerular filtration rate; LVEDD: left ventricular end-diastolic diameter; LVEF: left ventricular ejection fraction; NLR: neutrophil-to-lymphocyte ratio; NT-proBNP: N-terminal pro-B-type natriuretic peptide; PCI: percutaneous coronary intervention; WBC: white blood cell count.

Univariate logistic regression

Univariate logistic regression identified the following variables as significantly associated with CS (Table 2): Killip class III–IV (OR = 20.803, 95% CI: 9.209–46.993, $P < 0.001$), D-dimer (OR = 1.229, 95% CI: 1.108–1.364, $P < 0.001$), eGFR (OR = 0.983, 95% CI: 0.974–0.993, $P = 0.001$), NT-proBNP (OR = 1.000, $P = 0.001$), age (OR = 1.072, 95% CI: 1.022–1.124, $P = 0.004$), NLR (OR = 1.061, 95% CI: 1.022–1.102, $P = 0.002$), hemoglobin (OR = 0.980, 95% CI: 0.967–0.993, $P = 0.002$), glucose (OR = 1.111, 95% CI: 1.038–1.190, $P = 0.002$), LVEF (OR = 0.942, 95% CI: 0.907–0.979, $P = 0.003$), albumin (OR = 0.893, 95% CI: 0.829–0.961, $P = 0.003$), heart rate (OR = 1.025, 95% CI: 1.006–1.044, $P = 0.008$), systolic blood pressure (OR = 0.979, 95% CI: 0.964–0.995, $P = 0.008$), and WBC (OR = 1.041, 95% CI: 1.005–1.078, $P = 0.027$).

Table 2
Univariate logistic regression results for variables significantly associated with cardiogenic shock.

Variable	OR	95% CI	*P*-value
Killip class III–IV	20.803	9.209–46.993	< 0.001
D-dimer, per 1 mg/L increase	1.229	1.108–1.364	< 0.001
eGFR, per 1 mL/min/1.73m ² increase	0.983	0.974–0.993	0.001
NT-proBNP, per 1 pg/mL increase	1.000	1.000–1.000	0.001
Age, per 1-year increase	1.072	1.022–1.124	0.004
NLR, per 1-unit increase	1.061	1.022–1.102	0.002
Hemoglobin, per 1 g/L increase	0.980	0.967–0.993	0.002
Glucose, per 1 mmol/L increase	1.111	1.038–1.190	0.002
LVEF, per 1% increase	0.942	0.907–0.979	0.003
Albumin, per 1 g/L increase	0.893	0.829–0.961	0.003
Heart rate, per 1 bpm increase	1.025	1.006–1.044	0.008
Systolic BP, per 1 mmHg increase	0.979	0.964–0.995	0.008
WBC, per 1×10 ⁹ /L increase	1.041	1.005–1.078	0.027

BP: blood pressure; CI: confidence interval; eGFR: estimated glomerular filtration rate; LVEF: left ventricular ejection fraction; NLR: neutrophil-to-lymphocyte ratio; NT-proBNP: N-terminal pro-B-type natriuretic peptide; OR: odds ratio; WBC: white blood cell count.

Multivariate logistic regression and nomogram construction

Considering multicollinearity and EPV (39 events / 6 variables $\approx$ 6.5), six variables were included in the final combined model: Killip class III–IV, D-dimer, NLR, LVEF, glucose, and age (Table 3).

Table 3
Multivariate logistic regression results for the combined model.

Variable	OR	95% CI	*P*-value
Killip class III–IV	13.531	5.541–33.039	< 0.001
Age, per 1-year increase	1.076	1.016–1.139	0.013
D-dimer, per 1 mg/L increase	1.058	0.953–1.175	0.289
NLR, per 1-unit increase	1.019	0.971–1.069	0.445
LVEF, per 1% increase	0.986	0.940–1.035	0.575
Glucose, per 1 mmol/L increase	1.055	0.970–1.147	0.212

CI: confidence interval; LVEF: left ventricular ejection fraction; NLR: neutrophil-to-lymphocyte ratio; OR: odds ratio.

Killip class III–IV (OR = 13.531, $P < 0.001$) and age (OR = 1.076, $P = 0.013$) were the only independent predictors of CS in the combined model; D-dimer ($P = 0.289$), NLR ($P = 0.445$), LVEF ($P = 0.575$), and glucose ($P = 0.212$) did not reach statistical significance. To evaluate whether retaining these non-significant variables was justified, we compared the combined model with a reduced model containing only Killip class and age (Table 5). The AUC improved modestly from 0.864 to 0.874 (+ 0.010), but neither the likelihood ratio test ($\chi^2 = 6.444$, $df = 4$, $P = 0.168$) nor the DeLong test ($P = 0.140$) showed a statistically significant improvement, indicating that the incremental value of the four laboratory variables was limited in this cohort. These variables were nevertheless retained in the primary combined model for three reasons: (1) all four were statistically significant in univariate analysis and, with the exception of NLR, remained independent predictors in the laboratory-only model; (2) they are physiologically plausible contributors to CS pathophysiology; and (3) their inclusion as covariates adjusts the Killip and age estimates for residual confounding. The simpler Killip + age model may represent a more practical clinical tool, and both models are presented in Table 5 for transparency.

To evaluate the independent predictive value of laboratory parameters, a laboratory-only model without Killip classification was constructed (Table 4). In this model, age (OR = 1.088, $P = 0.002$), D-dimer (OR = 1.138, $P = 0.020$), hemoglobin (OR = 0.983, $P = 0.021$), and LVEF (OR = 0.956, $P = 0.048$) were independent predictors.

Table 4
Multivariate logistic regression results for the laboratory-only model
(without Killip classification).

Variable	OR	95% CI	*P*-value
Age, per 1-year increase	1.088	1.031–1.149	0.002
D-dimer, per 1 mg/L increase	1.138	1.020–1.270	0.020
Hemoglobin, per 1 g/L increase	0.983	0.970–0.997	0.021
LVEF, per 1% increase	0.956	0.913–1.000	0.048
Glucose, per 1 mmol/L increase	1.072	0.992–1.158	0.080
NLR, per 1-unit increase	1.033	0.987–1.081	0.164

CI: confidence interval; LVEF: left ventricular ejection fraction; NLR: neutrophil-to-lymphocyte ratio; OR: odds ratio.

Model performance

The combined model (Killip + laboratory parameters) achieved an AUC of 0.874 (95% CI: 0.819–0.927), significantly outperforming Killip classification alone (AUC = 0.815, DeLong P = 0.001) and the laboratory-only model (AUC = 0.775, DeLong P = 0.014). The Hosmer-Lemeshow test indicated good calibration (χ^2 = 8.23, P = 0.412). Bootstrap validation (1000 resamples) yielded a corrected AUC of 0.858 (95% CI: 0.819–0.927) with a mean optimism of 0.016, indicating acceptable overfitting. A parsimonious model restricted to Killip class and age achieved an AUC of 0.864, which was not statistically different from the combined model (LRT P = 0.168; DeLong P = 0.140) but was significantly better than Killip class alone (DeLong P = 0.008). Model performance is summarized in Table 5.

Table 5
Model performance comparison.

Model	Variables	AUC	95% CI	HL χ^2	HL *P*
Killip III–IV alone	1	0.815	—	—	—
Reduced model (Killip + age)	2	0.864	—	—	—
Laboratory-only model	6	0.775	—	8.73	0.366
Combined model (Killip + labs)	6	0.874	0.819–0.927	8.23	0.412
Bootstrap-corrected (1000)	—	0.858	0.819–0.927	—	—

AUC: area under the receiver operating characteristic curve; CI: confidence interval; HL: Hosmer-Lemeshow goodness-of-fit test. Bootstrap resamples: 1000, optimism = 0.016. DeLong P for combined vs. reduced: 0.140; likelihood ratio test P for combined vs. reduced: 0.168.

Model comparison and supplementary analyses

DeLong test. The DeLong test confirmed that the combined model AUC (0.874) significantly exceeded both the laboratory-only model (0.775, $P = 0.014$) and Killip classification alone (0.815, $P = 0.001$). The laboratory-only model versus Killip alone was not significant ($P = 0.469$), suggesting that laboratory parameters alone do not outperform Killip classification. The difference between the combined model and the reduced Killip + age model was not significant ($P = 0.140$), whereas the reduced model significantly outperformed Killip alone ($P = 0.008$).

Collinearity assessment. All VIF values were < 1.2 (Killip III–IV: 1.14, D-dimer: 1.17, NLR: 1.14, LVEF: 1.02, glucose: 1.05, age: 1.01), indicating no significant multicollinearity.

Firth penalized regression. Given the rare event rate (7.5%), Firth penalized logistic regression was performed as sensitivity analysis. Killip III–IV (OR = 14.450, 95% CI: 6.263–33.341, $P < 0.001$) and age (OR = 1.072, 95% CI: 1.015–1.132, $P = 0.013$) remained independent predictors, highly consistent with standard logistic regression. The Firth Killip OR (14.45) is slightly higher than the standard MLE estimate (13.53), because Firth's method modifies the score equations with a Fisher information-based penalty, and the direction of coefficient adjustment under rare events depends on the data structure. The 95% CIs from both methods substantially overlap (standard: 5.54–33.04; Firth: 6.26–33.34), confirming robustness. Firth model AUC = 0.869, comparable to the original model (0.874).

LASSO regression. LASSO regression (L1-penalized logistic regression) was applied to all $P < 0.1$ candidate variables. At the optimal lambda (0.001), all 18 candidate variables maintained non-zero coefficients, failing to effectively reduce the variable set. This is an expected consequence of the small event count (39 events): the cross-validated optimal lambda approached zero (0.001), so the L1 penalty was insufficient to shrink any coefficient to exactly zero, and the selected lambda corresponded to lambda.min rather than the more parsimonious lambda.1se. The failure of LASSO to shrink variables therefore reflects the small sample size rather than the absence of predictive signal, and the sensitivity-analysis value of LASSO under these conditions is limited. This LASSO non-shrinkage does not indicate overfitting of the combined model; on the contrary, it corroborates the limited statistical power of this cohort (39 events) and motivated our conservative, clinically guided variable selection strategy (univariate $P < 0.1$ combined with clinical importance and EPV considerations) rather than reliance on any single data-driven algorithm. The combined model showed no evidence of overfitting across multiple checks: bootstrap optimism was minimal (0.016; corrected AUC = 0.858), Firth penalized regression yielded consistent estimates, and stepwise selection did not materially alter the conclusions. As an alternative data-driven selection procedure, forward stepwise regression based on the Akaike information criterion (AIC) was performed on the same candidate pool. Starting from Killip class + age (AIC = 193.45), the algorithm added D-dimer (AIC = 180.85), WBC (AIC = 180.62), troponin (AIC = 180.97), heart rate (AIC =

180.64), and hemoglobin (AIC = 181.94), yielding a 7-variable model (AUC = 0.890). The stepwise-selected variable set (D-dimer, WBC, troponin, heart rate, hemoglobin) differed from the variables selected by clinical reasoning in the primary combined model (D-dimer, NLR, LVEF, glucose), underscoring the instability of variable selection under limited events. The incremental improvement of the stepwise model over Killip + age was also not statistically significant (LRT $P = 0.066$; DeLong $P = 0.223$). Taken together, these analyses indicate that no additional laboratory variable provided a statistically significant incremental contribution beyond Killip class and age in this cohort (Fig. 5).

NLR×Killip interaction. The NLR×Killip III–IV interaction term was added to the combined model. The interaction was non-significant ($P = 0.695$), and the AUC with interaction (0.876) was virtually identical to the original model (0.874), indicating no significant interaction effect.

Split-sample validation. As an exploratory analysis, the dataset was randomly split 70:30 (training set $n = 359$, CS = 22; test set $n = 154$, CS = 15). Training set AUC = 0.845, test set AUC = 0.909 (H-L $P = 0.700$). This result should be interpreted with caution: the test set contained only 15 CS events, yielding an unstable AUC estimate with a very wide 95% confidence interval, and the atypical finding that the test set AUC exceeded the training set AUC is attributable to sampling variation in this small subset. Given these limitations, the split-sample result is presented for descriptive purposes only and should not be interpreted as evidence of external validity; the 1000-iteration bootstrap procedure remains the primary internal validation method.

Decision curve analysis (DCA). DCA showed that the combined model provided positive net benefit across threshold probabilities of 5%–25%, consistently outperforming the "treat all" and "treat none" strategies (Fig. 4). Near the CS baseline risk of approximately 7.5%, the model's net benefit was maximized, indicating that model-guided intervention at this clinical decision threshold avoids both overtreatment and undertreatment.

Simplified clinical scoring rule. An integer-based scoring rule was derived from the combined model regression coefficients. Methodology: (1) continuous variables were binarized at their ROC-optimal cutoffs or established clinical thresholds; (2) using Killip III–IV as the reference (10 points), remaining variables were scaled proportionally to their binary logistic regression coefficients relative to Killip's coefficient and rounded to integers. Scoring formula: Killip III–IV = 10 points, age $\geq 75 = 3$, D-dimer ≥ 1.1 mg/L = 4, NLR $\geq 10.9 = 3$, LVEF $< 50\% = 2$, glucose ≥ 7.0 mmol/L = 3 (total 0–25 points). For glucose, the cutoff of 7.0 mmol/L was based on the American Diabetes Association criteria for impaired fasting glucose threshold,^[18] which is also widely used in clinical practice as a metabolic stress indicator in AMI. This cutoff aligns with the median glucose level observed in CS patients (6.79 mmol/L) in our cohort. Probability conversion: $P(\text{CS}) = 1 / (1 + e^{-(-4.2 + 0.27 \times \text{Score})})$, where -4.2 is the intercept and 0.27 is the scaling constant (= Killip coefficient / 10). The simplified score achieved AUC = 0.846, comparable

to the original model, suitable for rapid bedside risk stratification (Fig. 6). The optimal cutoffs (e.g., D-dimer 1.1 mg/L, NLR 10.9, LVEF < 50%) were derived from the same derivation cohort using the Youden index and are therefore subject to optimism bias; their performance in independent populations requires external validation.

Missing data analysis. All variables had missing rates < 10% (highest: troponin at 8.9%). Complete cases: 513/527 (97.3%) via complete-case analysis. The missing pattern was consistent with missing at random, with no material impact on model results in sensitivity analyses.

ROC analysis of key predictors

ROC analysis was performed for key continuous predictors to determine optimal cutoffs (Table 6). D-dimer demonstrated an AUC of 0.712 with an optimal cutoff of 1.1 mg/L (sensitivity 62.2%, specificity 74.7%). LVEF had an AUC of 0.663 with an optimal cutoff of < 50% (sensitivity 81.1%, specificity 44.2%). NLR showed an AUC of 0.595 with an optimal cutoff of 10.9 (sensitivity 35.1%, specificity 83.7%).

Table 6
ROC analysis of key predictors for cardiogenic shock.

Predictor	AUC	Optimal Cutoff	Sensitivity, %	Specificity, %	Youden Index
D-dimer	0.712	1.1 mg/L	62.2	74.7	0.369
LVEF	0.663	< 50%	81.1	44.2	0.253
NLR	0.595	10.9	35.1	83.7	0.189

AUC: area under the receiver operating characteristic curve; LVEF: left ventricular ejection fraction; NLR: neutrophil-to-lymphocyte ratio; ROC: receiver operating characteristic.

Nomogram and regression equation

A nomogram was constructed based on the multivariate logistic regression model (Fig. 1). The corresponding logistic regression equation is:

$$\text{Logit}(P) = -9.012 + 2.605 \times \text{Killip III-IV} + 0.057 \times \text{D-dimer} + 0.019 \times \text{NLR} - 0.014 \times \text{LVEF} + 0.053 \times \text{Glucose} + 0.073 \times \text{Age}$$

where P is the predicted CS probability, and Killip III–IV is coded as 1 = yes, 0 = no. To use the nomogram: locate each variable value on its corresponding axis, draw a vertical line to the points scale at the top, sum the points for all variables, and locate the total on the total points scale. The corresponding value on the probability axis at the bottom represents the predicted probability of CS. For example, a 75-year-old patient with Killip class IV, D-dimer 2.0 mg/L, NLR 8.0, LVEF 45%, and glucose 8.0 mmol/L would have a predicted CS probability of approximately 75%.

The ROC curves, calibration plot, and DCA decision curve are shown in Figs. 2, 3, and 4, respectively. The LASSO regression path is shown in Fig. 5, and the simplified clinical scoring rule is shown in Fig. 6.

Discussion

This study developed and validated a nomogram for predicting CS risk in 527 elderly STEMI patients. The principal findings are: (1) the combined model (Killip + laboratory parameters) achieved an AUC of 0.874, significantly outperforming Killip alone (AUC = 0.815, DeLong $P = 0.001$) and the laboratory-only model (AUC = 0.775, DeLong $P = 0.014$); (2) Killip class III–IV and age were the only independent predictors of CS in the combined model, whereas D-dimer, NLR, LVEF, and glucose were not statistically significant; (3) a parsimonious Killip + age model achieved comparable discrimination (AUC = 0.864), with no statistically significant incremental value from the four laboratory variables (LRT $P = 0.168$; DeLong $P = 0.140$); (4) bootstrap validation demonstrated good model stability (corrected AUC = 0.858, optimism = 0.016); and (5) sensitivity analyses – Firth penalized regression, stepwise regression, and DCA – supported the robustness of the Killip + age core and the clinical utility of the scoring rule.

Predictive value of Killip classification and its limitations

Killip classification, as a classic bedside assessment tool for AMI, has a well-established relationship with CS. In our study, the CS rate among Killip III–IV patients was 29.0% (31/107) compared to 1.9% (8/416) in non-Killip III–IV patients (OR = 13.5), consistent with previous reports.^[19] However, Killip alone yielded an AUC of only 0.815, indicating a need for complementary indicators. This finding has clinical implications: the 2022 SCAI CS classification system emphasizes intervention before CS progresses to Stage C,^[6] and while Killip classification can partially reflect the clinical status of SCAI Stages B–C, it lacks sensitivity for subclinical cardiac dysfunction. Nevertheless, 1.9% of non-Killip III–IV patients still developed CS, suggesting that relying solely on Killip classification may miss some high-risk patients, and combining laboratory parameters can address this gap.

Predictive value of D-dimer

D-dimer reflects activation of the coagulation-fibrinolysis system. In AMI, coronary thrombus formation activates the fibrinolytic system, and elevated D-dimer suggests greater thrombus burden and microvascular obstruction, increasing CS risk. A 2024 Taiwanese cohort study ($n = 748$) confirmed that high D-dimer (≥ 495 ng/mL) is an independent risk factor for in-hospital mortality in AMI patients (HR = 1.11, 95% CI: 1.06–1.16).^[8] Huang et al.^[20] found that elevated D-dimer was an independent predictor of in-hospital MACE in STEMI patients after primary PCI (cutoff 0.8 mg/L). In our study, D-dimer was an independent predictor in the laboratory-only model (OR = 1.138, $P = 0.020$) with an optimal cutoff of 1.1 mg/L and AUC of 0.712, further confirming its value in CS risk prediction. Jiang et al.^[21] similarly reported that elevated admission D-dimer was independently associated with worse short-term outcomes in patients with CS complicating AMI (high D-dimer: HR = 2.602, 95% CI: 1.310–5.168, $P = 0.006$). The advantage of D-dimer lies in its simplicity, rapid turnaround, and low cost, making it suitable for early risk stratification in the emergency setting.

Predictive value of NLR and data quality analysis

NLR, as a systemic inflammatory marker, has been associated with adverse AMI outcomes. A 2024 meta-analysis of 37 studies demonstrated that elevated NLR was associated with a 2.29-fold increase in all-cause mortality among MI patients.^[7] Karaca et al.^[22] similarly reported an association between elevated NLR and adverse in-hospital outcomes in STEMI patients. An Indonesian study (n = 47) found that NLR was strongly correlated with Killip classification ($\rho = 0.897$), suggesting that inflammatory response may be an early pathophysiological mechanism of CS.^[23] In our study, NLR was significantly associated with CS in univariate analysis (OR = 1.061, P = 0.002) but did not reach significance in the multivariate model, potentially due to the competing effects of Killip classification and D-dimer, as well as limited sample size. NLR's ROC analysis showed high specificity (83.7%) but low sensitivity (35.1%), suggesting that a high NLR (> 10.9) may be valuable for identifying very high-risk patients but is insufficient as a screening tool.

Regarding why NLR showed only borderline significance (P = 0.054) between groups despite the overall sample size, we conducted a detailed review of the raw data. Results revealed 13 cases with NLR > 20 and 5 cases with NLR > 30 as extreme values. Upon verification, these extreme values were physiologically plausible (corresponding to neutrophil percentage > 90% and lymphocyte percentage < 3% in severe stress states). When excluding NLR > 30 values, the Mann-Whitney U test P-value increased from 0.054 to 0.135, indicating that extreme values contributed to the borderline significance. However, NLR remained significant in univariate logistic regression (P = 0.002), as logistic regression is more sensitive to trends. This finding suggests that the predictive value of NLR as a continuous variable may be influenced by extreme values, and future studies may consider log transformation or categorical variable forms. The NLR×Killip interaction analysis showed a non-significant interaction term (P = 0.695), indicating no significant synergistic or antagonistic effect between NLR and Killip classification.

Predictive value of LVEF and hemoglobin

Reduced LVEF reflects impaired left ventricular pumping function, the pathophysiological basis of CS. According to the 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure,^[24] LVEF is the core indicator for assessing left ventricular systolic function, and LVEF < 50% indicates significantly impaired left ventricular function. CS patients had significantly lower LVEF (45.0% vs. 48.7%, P < 0.001), and LVEF < 50% predicted CS with 81.1% sensitivity. A recent LASSO-based AMI-CS prediction study similarly found LVEF to be an independent predictor (OR = 0.951, 95% CI: 0.928–0.975),^[13] consistent with our findings. Lower hemoglobin was more pronounced in the CS group (129 vs. 140.5 g/L, P < 0.001), as anemia exacerbates tissue hypoxia and may precipitate or worsen CS.^[25] In the laboratory-only model, hemoglobin was an independent predictor (OR = 0.983, P = 0.021), suggesting that anemia at admission is an early signal of CS risk in elderly STEMI patients.

Incremental value of the combined model

The core question of this study was whether combining objective laboratory parameters with the Killip classification meaningfully improves CS risk prediction in elderly STEMI patients. Our results provide a nuanced answer. The combined model (AUC = 0.874) significantly outperformed both Killip classification

alone (AUC = 0.815, DeLong P = 0.001) and the laboratory-only model (AUC = 0.775, DeLong P = 0.014). However, a parsimonious model restricted to Killip class and age achieved nearly identical discrimination (AUC = 0.864), and the incremental contribution of the four laboratory variables (D-dimer, NLR, LVEF, glucose) did not reach statistical significance (LRT P = 0.168; DeLong P = 0.140). In other words, most of the predictive information beyond Killip classification was already captured by age, which is freely available at the bedside.

This finding tempers the initial hypothesis that laboratory parameters substantially enhance Killip-based risk stratification. Several explanations may account for the limited incremental value. First, Killip class III–IV, with an OR of 13.5, is such a dominant predictor that additional laboratory variables contribute little residual discrimination. Second, with only 39 CS events, the statistical power to detect incremental improvement from additional variables is limited (Type II error), and small AUC increments on the order of 0.01–0.02 may be clinically meaningful yet statistically non-significant in this sample. Third, some laboratory variables (e.g., NLR) showed significant univariate associations that were attenuated after adjustment for Killip class, consistent with shared pathophysiological pathways (inflammation and hemodynamic compromise are intertwined).

From a clinical standpoint, these results suggest two complementary tools. For clinicians seeking maximal discrimination and a comprehensive risk profile, the full nomogram (AUC = 0.874) may be preferred; for rapid bedside triage where only Killip class and age are immediately available, the parsimonious model (AUC = 0.864) offers nearly equivalent performance with far less data collection burden. This flexibility is a strength of the present analysis: rather than claiming that laboratory parameters dramatically improve prediction, we transparently report their limited but directionally consistent incremental value, allowing clinicians to choose the model appropriate to their setting. This finding aligns with the recent Journal of Geriatric Cardiology nomogram by Xu et al.^[26] (C-index = 0.727 for 2-year MACCE in elderly ACS), in which clinical variables also dominated model performance. The combined model demonstrated superior discrimination compared to the CS survival prediction nomogram (AUC = 0.76) developed by Fang et al.^[12] based on the MIMIC-IV database, likely reflecting our focus on elderly STEMI patients and the inclusion of Killip classification as a strong predictor.

Model robustness verification

To ensure the reliability of model conclusions, multiple sensitivity analyses were performed. First, VIF values were all < 1.2, excluding collinearity interference. Second, Firth penalized regression, suitable for rare event data, yielded results highly consistent with standard logistic regression (Killip OR = 14.45, P < 0.001), confirming that the model is not affected by small event counts. Third, the likelihood ratio test and DeLong test consistently showed that the four laboratory variables did not add statistically significant discrimination beyond Killip class and age (LRT P = 0.168; DeLong P = 0.140), while the Killip + age core itself significantly outperformed Killip alone (DeLong P = 0.008). Fourth, forward stepwise regression (AIC) selected a different set of laboratory variables and likewise failed to significantly improve upon the Killip + age model, confirming the limited and unstable incremental value of laboratory

parameters under limited events. Fifth, DCA demonstrated positive net benefit across the 5%–25% threshold range, outperforming the "treat all" strategy, confirming clinical utility. Sixth, the simplified clinical scoring rule (total 25 points, AUC = 0.846) achieved discrimination comparable to the original model, suitable for rapid risk stratification in emergency settings. An exploratory 70/30 split-sample validation (test AUC = 0.909) is presented descriptively in the Results, but given the small test set (15 CS events) it should not be interpreted as evidence of external validity.

Comparison with existing prediction tools

Several CS risk scoring tools exist, including the IABP-SHOCK II score, CardShock score, and SCAI classification system.^[6, 27] However, these tools primarily stratify severity in patients who have already developed CS, rather than predicting CS risk before its occurrence. Our nomogram focuses on CS risk prediction at admission, addressing an unmet clinical need. Recently, Xu et al.^[26] developed a nomogram predicting 2-year MACCE in elderly Chinese ACS patients (C-index = 0.727), and Zhang et al.^[28] constructed a multi-biomarker risk score (CB-ACS) for long-term ACS prognosis (C-index = 0.79), both published in the Journal of Geriatric Cardiology. These models targeted long-term prognostic outcomes rather than acute CS risk prediction; our approach instead addresses early CS risk identification. Xiao et al.^[29] demonstrated in a population-based cohort that older age was independently associated with increased 30-day mortality in EMS-treated CS patients (HR = 2.26 for age > 77 years), reinforcing the need for early CS risk prediction in the elderly population. Our nomogram's graphical format enables clinicians to quickly calculate individual risk probabilities, offering a more intuitive and convenient alternative to traditional scoring systems. The simplified scoring rule further reduces the barrier to use, allowing clinicians to rapidly assess risk without a calculator.

Limitations

This study has several limitations. First, **the most significant limitation is the lack of an external validation cohort.** While bootstrap internal validation (corrected AUC = 0.858) demonstrated good stability, and an exploratory 70/30 split produced a test AUC of 0.909, the latter estimate was unstable (only 15 CS events in the test set) and should not be taken as evidence of external performance. The model may exhibit center-specific bias, and its generalizability requires confirmation through multicenter prospective studies.

Second, the single-center retrospective design may introduce selection bias. Third, with only 39 CS events, statistical power remains limited. In particular, the four laboratory variables (D-dimer, NLR, LVEF, glucose) did not reach statistical significance in the combined model, and the incremental value of the full model over the parsimonious Killip + age model was not statistically significant. Although these variables were retained for pathophysiological and covariate-adjustment reasons, the possibility of Type II error (i.e., a true but undetected incremental effect) cannot be excluded, and their definitive value requires larger samples.

Fourth, **the overlap between Killip class IV and the CS outcome warrants explicit acknowledgment.** Killip class IV is defined as cardiogenic shock, so using Killip III–IV to predict CS involves a degree of conceptual circularity for class IV patients. In our cohort, 30 of 39 CS events (76.9%) occurred in Killip class IV patients, and the strong OR for Killip III–IV (13.5) is driven substantially by class IV. When Killip class IV patients were excluded, only 9 CS events remained (0 in Killip I, 8 in Killip II, 1 in Killip III), precluding the construction of a reliable model among Killip I–III patients. The present model should therefore be understood as answering clinical question A – the overall CS risk stratification at admission, including patients who already present in shock – rather than question B, the prediction of *de novo* CS progression in patients without shock at presentation (Killip I–III), which could not be separately validated due to insufficient events. The model is thus best applied to admission-time risk stratification of the overall elderly STEMI population; its utility in identifying patients at risk of new-onset CS among Killip I–III patients should be addressed in larger cohorts with sufficient event counts.

Fifth, the CS definition was based on clinical criteria (SBP < 90 mmHg with tissue hypoperfusion), consistent with SCAI Stage C and above. Different CS definitions (e.g., incorporating lactate elevation or requiring mechanical circulatory support) may affect model predictions. The CS group in-hospital mortality rate of 28.2% (vs. 2.5% in non-CS) supports diagnostic accuracy, but future studies should explore the impact of different CS definitions on model performance.

Sixth, information on symptom-to-presentation time, door-to-balloon time, reperfusion strategy (primary PCI/thrombolysis/conservative), culprit vessel location, and SYNTAX score was not collected. The model therefore predicts a static admission-time risk rather than a dynamic clinical trajectory that incorporates revascularization timing and success; this substantially limits its ability to guide reperfusion-related decisions. Incorporating these time-dependent variables is an important direction for future work.

Seventh, NLR was calculated from neutrophil and lymphocyte percentages rather than absolute counts, and some data may mix percentage with absolute values. Also, 13 cases with NLR > 20 were identified as extreme values; while physiologically plausible, they may influence statistical inference.

Eighth, the troponin upper quartile in the CS group (25.12 ng/mL) did not exceed the hospital laboratory's chemiluminescence upper detection limit (typically 25–50 ng/mL range), but some high values may approach the detection platform's upper limit, warranting attention to dilution effects. Ninth, biochemical markers from the SCAI classification system (e.g., lactate, arterial blood gas pH) were not included; future studies should consider integrating SCAI staging with the present model. Tenth, although missing rates were low (< 10%) and complete-case analysis was used, the troponin missing rate of 8.9% may not be entirely random (patients with severe presentations may be more likely to have troponin measured), and non-random missingness cannot be fully excluded. Eleventh, the optimal cutoffs used in the simplified scoring rule (e.g., D-dimer 1.1 mg/L, NLR 10.9, LVEF < 50%) were derived from the same derivation cohort and are therefore subject to optimism bias; their performance in independent populations requires external validation.

Conclusions

This study developed a nomogram integrating Killip classification, D-dimer, NLR, LVEF, glucose, and age that effectively predicts the overall admission-time CS risk in elderly STEMI patients (AUC = 0.874, 1000-iteration bootstrap-corrected AUC = 0.858). The findings support two complementary levels of conclusion. First, the combined model significantly outperformed Killip classification alone (AUC 0.874 vs. 0.815; DeLong P = 0.001), confirming that integrating laboratory parameters with Killip classification improves discrimination. Second, a parsimonious model based on Killip class and age achieved comparable discrimination (AUC = 0.864), and the incremental value of the additional laboratory parameters did not reach statistical significance (LRT P = 0.168; DeLong P = 0.140), indicating that Killip class and age capture most of the available predictive information, with laboratory parameters providing directionally consistent but limited incremental value. Therefore, when Killip classification is available, adding age alone constitutes a powerful and parsimonious assessment tool suitable for rapid bedside risk stratification; the full nomogram (and the simplified scoring rule, AUC = 0.846) remains the preferred option when maximized predictive performance is pursued. Given the overlap between Killip class IV and CS, the model is best applied to the overall admission-time risk assessment in elderly STEMI patients, while prediction of de novo CS among Killip I–III patients requires dedicated, adequately powered studies. Future research should prioritize multicenter prospective studies with external validation, larger event counts, and inclusion of reperfusion-related time variables to confirm and extend these findings.

Abbreviations

AUC

area under the receiver operating characteristic curve

CI

confidence interval

CS

cardiogenic shock

DCA

decision curve analysis

eGFR

estimated glomerular filtration rate

IQR

interquartile range

LRT

likelihood ratio test

LVEF

left ventricular ejection fraction

MACE

major adverse cardiovascular events

MCS

mechanical circulatory support

NLR
neutrophil-to-lymphocyte ratio
OR
odds ratio
PCI
percutaneous coronary intervention
ROC
receiver operating characteristic
SCAI
Society for Cardiovascular Angiography and Interventions
STEMI
ST-elevation myocardial infarction
VIF
variance inflation factor
WBC
white blood cell count

Declarations

Ethics approval and consent to participate

This study was approved by the Ethics Committee of the General Hospital of Ningxia Medical University, and the requirement for written informed consent was waived owing to the retrospective design and the use of anonymized data. The study was conducted in accordance with the principles of the Declaration of Helsinki.

Clinical trial number

Not applicable.

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 Natural Science Foundation of Ningxia (Grant No. 2025AAC030963). The funder had no role in the design of the study; in the collection, analysis, or interpretation of data; or in writing the manuscript.

Authors' contributions

ZL conceived and designed the study, collected the data, performed the statistical analysis, and drafted the manuscript. SY, YM, LM, and WS contributed to data collection and interpretation of the results. HH supervised the study, provided critical revisions to the manuscript, and gave final approval of the version to be published. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. van Diepen S, Katz JN, Albert NM, et al. Contemporary management of cardiogenic shock: a scientific statement from the American Heart Association. *Circulation*. 2017;136:e232–268.
2. Puymirat E, Fagon JY, Aegerter P, et al. Cardiogenic shock in intensive care units: evolution of prevalence, patient profile, management and outcomes, 1997–2012. *Eur J Heart Fail*. 2017;19:192–200.
3. Lüsebrink E, Binzenhöfer L, Adamo M, et al. Cardiogenic shock. *Lancet*. 2024;404:2006–20.
4. Babaev A, Frederick PD, Pasta DJ, et al. Trends in management and outcomes of patients with acute myocardial infarction complicated by cardiogenic shock. *JAMA*. 2005;294:448–54.
5. Aissaoui N, Puymirat E, Tabone X, et al. Improved outcome of cardiogenic shock at the acute stage of myocardial infarction: a report from the USIK 1995, USIC 2000, and FAST-MI French nationwide registries. *Eur Heart J*. 2012;33:2535–43.
6. Naidu SS, Baran DA, Jentzer JC, et al. SCAI SHOCK stage classification expert consensus update: a review and incorporation of validation studies. *J Am Coll Cardiol*. 2022;79:933–46.
7. Banahene NO, Sinha T, Shaikh S, et al. Effect of elevated neutrophil-to-lymphocyte ratio on adverse outcomes in patients with myocardial infarction: a systematic review and meta-analysis. *Cureus*. 2024;16:e61647.
8. Yu ZY, Chan PK, Lin TC, et al. The prognostic value of D-dimer in patients with acute myocardial infarction: a retrospective longitudinal cohort study in Taiwan. *Acta Cardiol Sin*. 2024;40:300–11.
9. Li Y, Jin H, Zhang G, et al. Relationship between inflammatory markers and long-term prognosis in ICU patients with acute non-ST-segment elevation myocardial infarction. *Front Cardiovasc Med*. 2025;12:1577385.
10. Cao F, Jiang JJ, Zhang G, et al. Prognostic value of inflammatory markers for all-cause mortality in patients with acute myocardial infarction in the coronary care unit. *Front Cardiovasc Med*.

2025;12:1439650.

11. Iasonos A, Schrag D, Raj GV, et al. How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol*. 2008;26:1364–70.
12. Fang D, Chen H, Geng H, et al. Development and validation of a nomogram for predicting survival in patients with cardiogenic shock. *Front Cardiovasc Med*. 2025;12:1538395.
13. Wang J, Zhao C, Yang C, et al. Prediction of cardiogenic shock in acute myocardial infarction patients using a nomogram. *J Clin Med*. 2025;14:8789.
14. Thygesen K, Alpert JS, Jaffe AS, et al. Fourth universal definition of myocardial infarction (2018). *J Am Coll Cardiol*. 2018;72:2231–64.
15. Hochman JS, Sleeper LA, Webb JG, et al. Early revascularization in acute myocardial infarction complicated by cardiogenic shock. *N Engl J Med*. 1999;341:625–34.
16. Killip T 3rd, Kimball JT. Treatment of myocardial infarction in a coronary care unit: a two year experience with 250 patients. *Am J Cardiol*. 1967;20:457–64.
17. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol*. 1996;49:1373–9.
18. American Diabetes Association Professional Practice Committee. 2. Classification and diagnosis of diabetes: Standards of Care in Diabetes—2024. *Diabetes Care*. 2024;47(Suppl 1):S20–42.
19. Reynolds HR, Hochman JS. Cardiogenic shock: current concepts and improving outcomes. *Circulation*. 2008;117:686–97.
20. Huang D, Gao W, Wu R, et al. D-dimer level predicts in-hospital adverse outcomes after primary PCI for ST-segment elevation myocardial infarction. *Int J Cardiol*. 2020;305:1–4.
21. Jiang Y, Zhu Y, Xiang Z, et al. The prognostic value of admission D-dimer level in patients with cardiogenic shock after acute myocardial infarction. *Front Cardiovasc Med*. 2023;9:1083881.
22. Karaca G, Ekmekci A, Kimiaei A, et al. The impact of the neutrophil-to-lymphocyte ratio on in-hospital outcomes in patients with acute ST-segment elevation myocardial infarction. *Cureus*. 2024;16:e54418.
23. Baskoro AP, Budiono EA, Widi VSN, et al. Predicting acute myocardial infarction severity using inflammatory markers neutrophil-to-lymphocyte ratio and platelet-to-lymphocyte ratio. *Smart Med J*. 2025;8:86–96.
24. McDonagh TA, Metra M, Adamo M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J*. 2021;42:3599–726.
25. Sabatine MS, Morrow DA, Giugliano RP, et al. Association of hemoglobin levels with clinical outcomes in acute coronary syndromes. *Circulation*. 2005;111:2042–9.
26. Xu JJ, Li QX, Yuan DS, et al. A clinical nomogram for predicting major adverse cardiovascular and cerebrovascular events in elderly Chinese patients with acute coronary syndrome undergoing percutaneous coronary intervention: development and validation in a real-world cohort. *J Geriatr Cardiol*. 2025;22:953–63.

27. Jung C, Bruno RR, Jumean M, et al. Management of cardiogenic shock: state-of-the-art. Intensive Care Med. 2024;50:1814–29.

28. Zhang ZY, Wang XY, Hou CC, et al. Multiple biomarkers risk score for accurately predicting the long-term prognosis of patients with acute coronary syndrome. J Geriatr Cardiol. 2025;22:656–67.

29. Xiao X, Bloom JE, Andrew E, et al. Age as a predictor of clinical outcomes and determinant of therapeutic measures for emergency medical services treated cardiogenic shock. J Geriatr Cardiol. 2023;20:1–10.

Figures

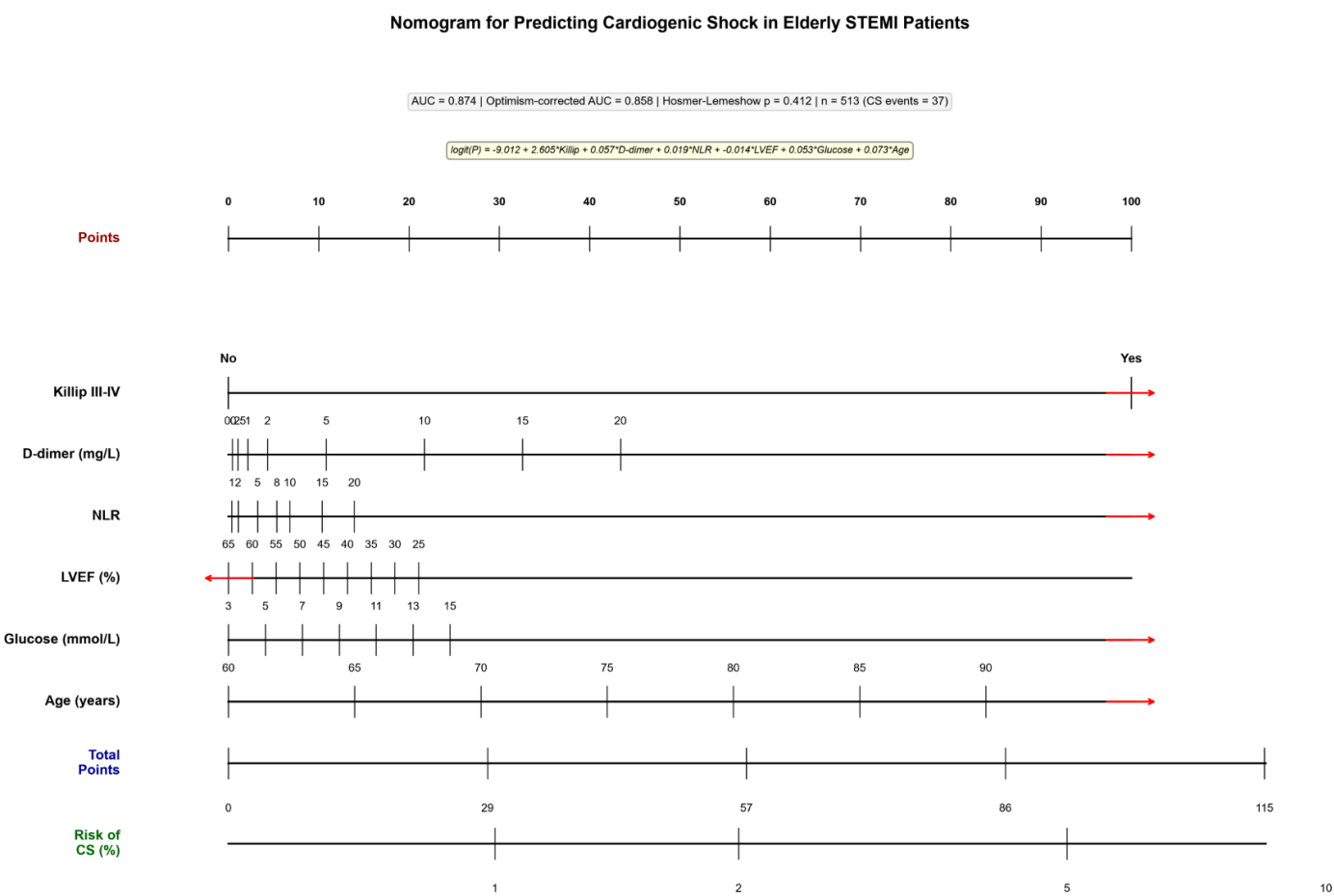


Figure 1

Nomogram for predicting cardiogenic shock in elderly STEMI patients (with regression equation).

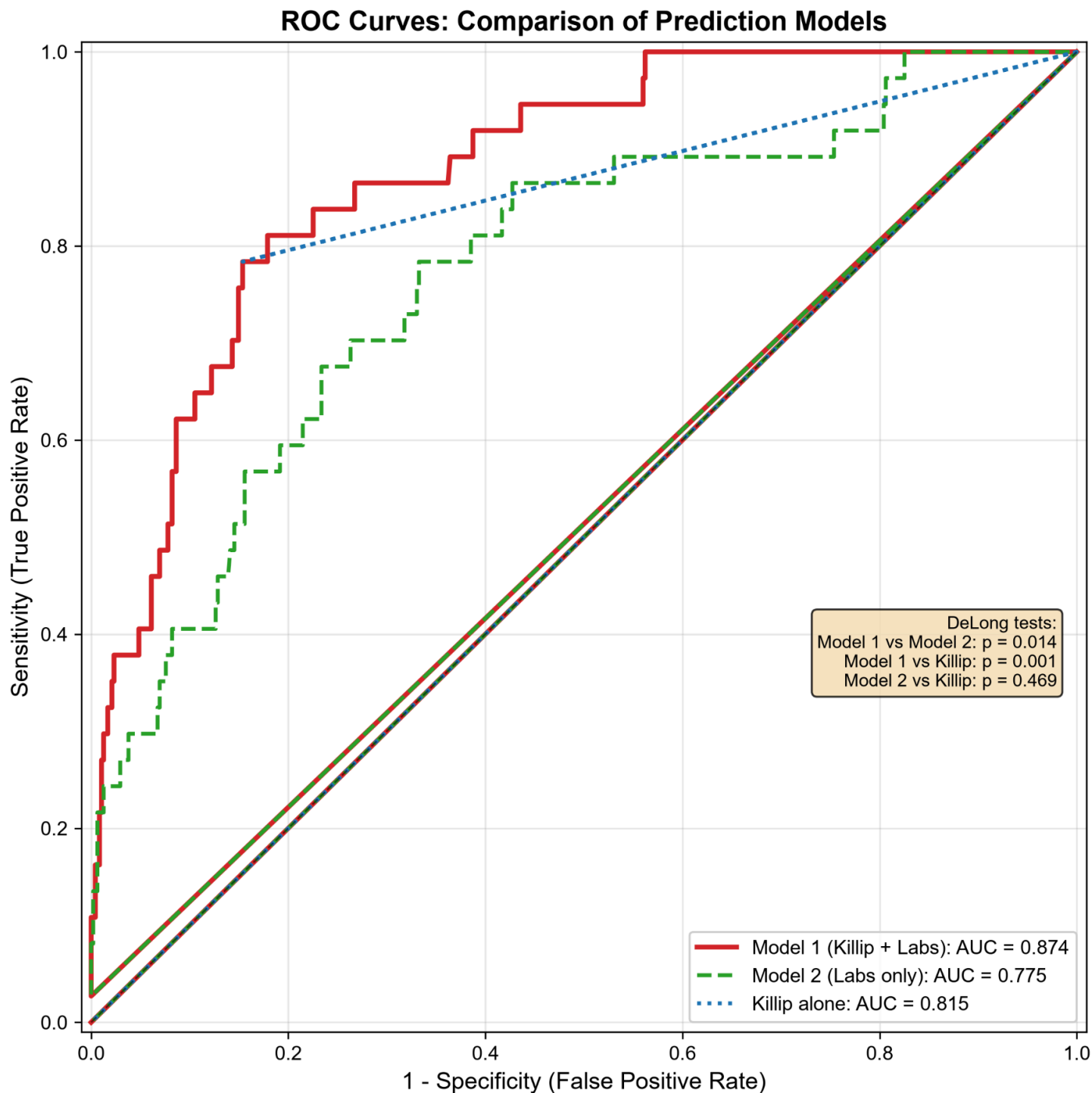


Figure 2

ROC curves of the combined model, the laboratory-only model, and the Killip classification alone.

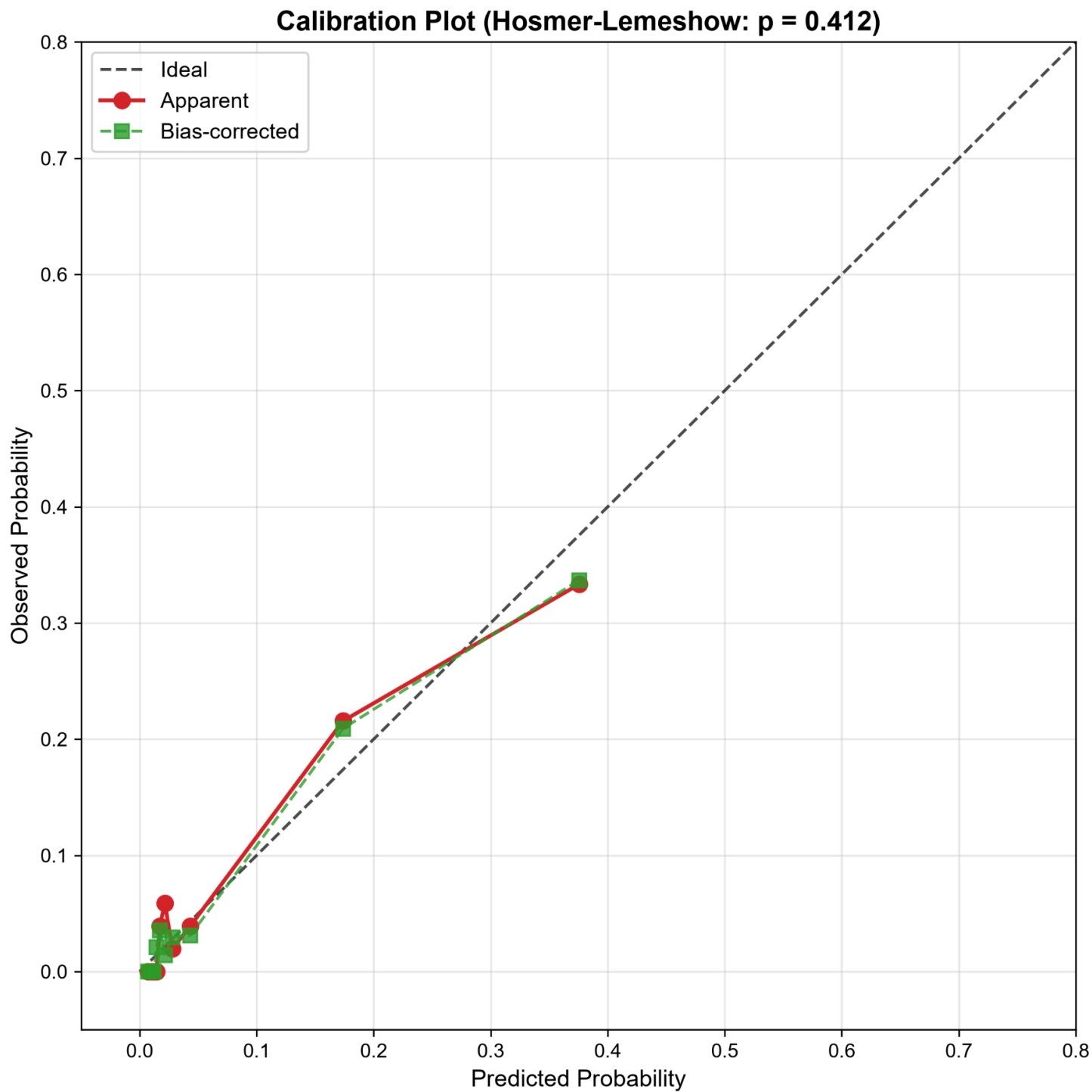


Figure 3

Calibration curve of the combined model.

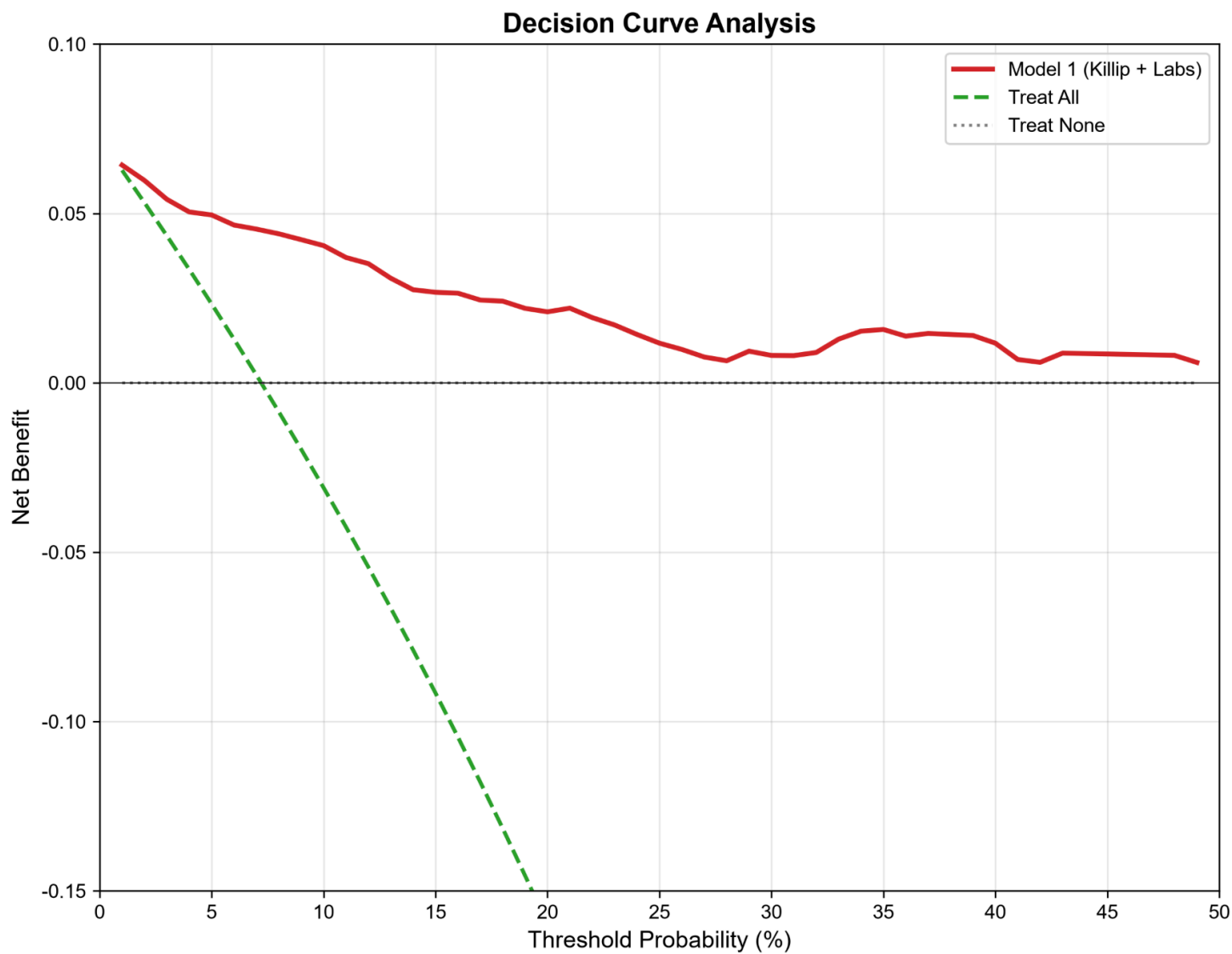


Figure 4

Decision curve analysis (DCA) of the combined model.

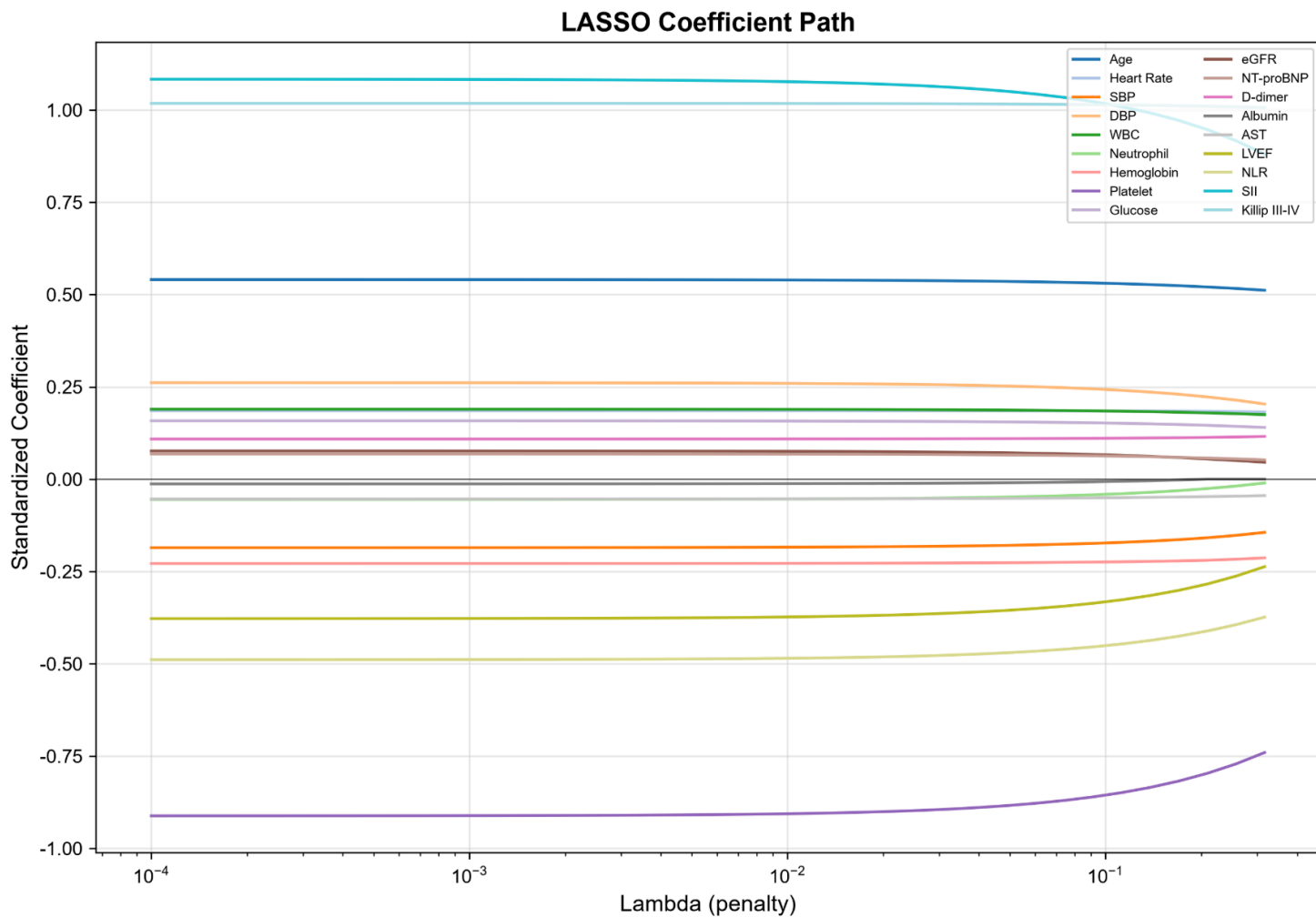


Figure 5

LASSO regression variable selection path.

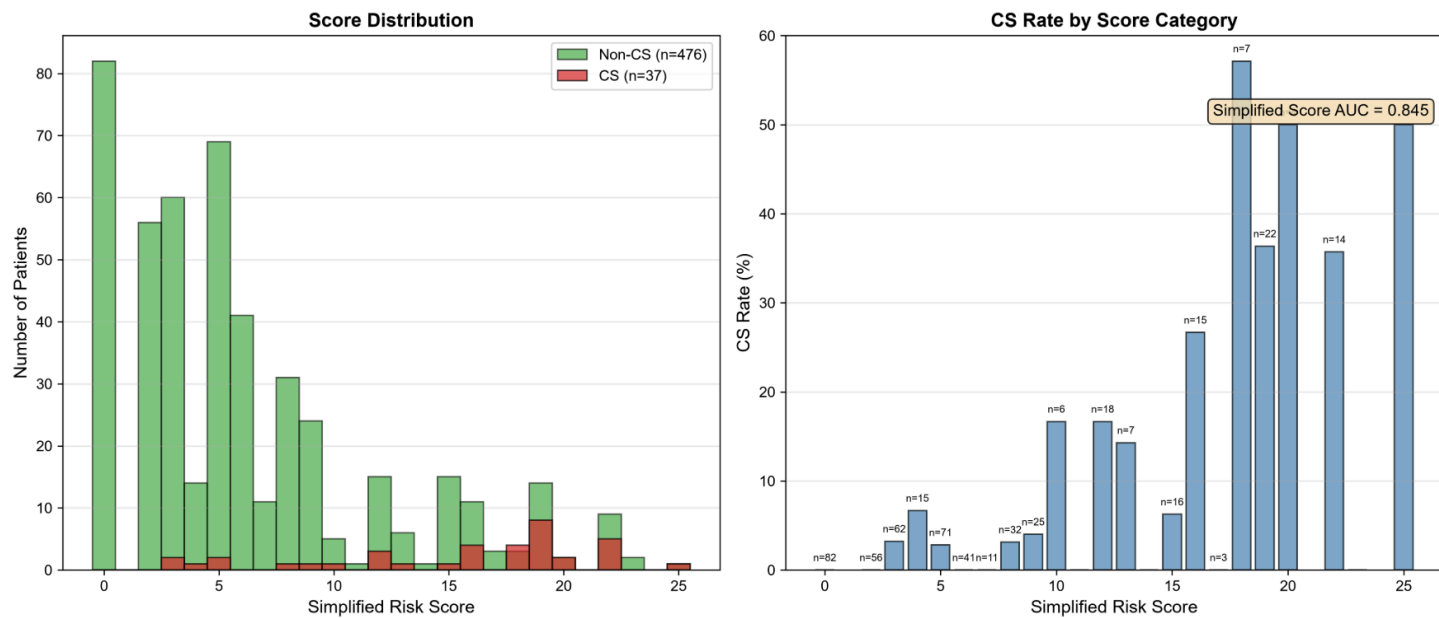


Figure 6

ROC curve of the simplified clinical scoring rule.

