e-ISSN: 0976-822X, p-ISSN:2961-6042

Available online on http://www.ijcpr.com/

International Journal of Current Pharmaceutical Review and Research 2026; 18(6); 369-372

Original Research Article

Prediction Models for Major Adverse Cardiovascular Events Following ST-Segment Elevation Myocardial Infarction and Subgroup-Specific Performance

Ravi Prakshbhai Patel[1], Vaidya Honeyben Chandrakant[2], Kapil Khanna[3]

[1]MBBS, Batumi Shota Rustaveli State University Georgia
[2]MBBS, Grodno state Medical University
[3]MD, PGDCCP, Consultant Cardio-Physician, Yashoda Medicity Hospital, Indirapuram, Ghaziabad, Uttar Pradesh, India

Received: 04-04-2026 / Revised: 04-05-2026 / Accepted: 05-06-2026
Corresponding Author: Dr. Kapil Khanna
Conflict of interest: Nil

Abstract:

Background: Despite primary percutaneous coronary intervention (PCI) and guideline-directed therapy, patients who survive ST-segment elevation myocardial infarction (STEMI) continue to be at risk of major adverse cardiovascular events (MACE). Personalized follow-up may be aided by prediction models, but there is limited reporting on model performance across subgroups.

Methods: A retrospective cohort of 680 adults with STEMI who underwent primary or rescue PCI was analysed. The primary outcome was 12-month major adverse cardiovascular events, defined as all-cause death, recurrent myocardial infarction, stroke, target-vessel revascularization or heart failure hospitalization. A total of 42 baseline clinical, angiographic, laboratory and treatment parameters were used to develop logistic regression, random forest and gradient boosting models. Model discrimination and calibration and subgroup performance were assessed.

Results: 118 patients (17.4%) had 12-month MACE. MACE patients were older than non-MACE patients, had more diabetes, chronic kidney disease, anterior wall infarction, multivessel disease, higher Killip class and lower left ventricular ejection fraction. Gradient boosting had the highest discrimination (AUC 0.84; 95% CI 0.79-0.89), followed by random forest (AUC 0.82) and logistic regression (AUC 0.78). Gradient boosting had a sensitivity of 78.0%, and a specificity of 77.4%. The best performance was observed in non-diabetic patients (AUC 0.86) and the worst in chronic kidney disease (AUC 0.77). Killip class, LVEF, creatinine, diabetes, symptom-to-balloon time, multivessel disease, final TIMI flow and discharge statin/non-adherence risk were important predictors.

Conclusion: Gradient boosting model outperformed MACE prediction following STEMI and demonstrated clinically significant differences between subgroups. Prediction tools should be reported to be calibrated and perform well in subgroups before clinical use.

Keywords: STEMI; Major Adverse Cardiovascular Events; Prediction Model; Machine Learning; Subgroup Analysis; Percutaneous Coronary Intervention.

DOI: 10.25258/ijcpr.18.6.63

Introduction

ST-segment elevation myocardial infarction is a time-dependent emergency and early reperfusion decreases infarct size, maintains ventricular function and increases survival [1]. Where available, primary PCI is the preferred reperfusion strategy, but there is still significant residual risk due to microvascular obstruction, recurrent ischemia, heart failure, arrhythmia, and incomplete revascularization and comorbidity burden [2].

Patients do not have an even distribution of major adverse cardiovascular events following STEMI. Prognosis is modified by age, diabetes, renal dysfunction, Killip class, infarct location, left ventricular ejection fraction, angiographic complexity and treatment adherence [3]. Traditional risk scores are helpful, but many were created prior to modern PCI, powerful antiplatelet agents, and high-intensity statins. It is therefore necessary to re-evaluate the modern prediction modelling with the latest data sets and to validate it carefully [4,5].

Non-linear interactions between clinical, laboratory and angiographic variables can be modeled using machine learning approaches. A recent study specifically assessed prediction models after STEMI and emphasized the importance of assessing the performance of these models in subgroups as risk

factors may differ in diabetics, elderly patients, women and patients with renal dysfunction [6]. Other acute coronary syndrome studies indicate that machine learning can provide reasonable discrimination, though transportability and interpretability are still issues [7,8].

The primary goal of the present study was to create prediction models for 12-month MACE after STEMI treated with PCI, to compare logistic regression with tree-based machine learning models, and to evaluate the performance of the models in subgroups. The goal was to not only find the best performing model, but also to see if the model was clinically valid in high-risk subgroups.

Materials and Methods

Study Design and Population: This was a retrospective cohort study of adults who were admitted with STEMI and received primary or rescue PCI over a three-year period. STEMI was characterized by the presence of ischemic symptoms and persistent ST-segment elevation or new left bundle branch block, along with biomarker evidence of myocardial necrosis. Patients who died prior to angiography, were treated with thrombolysis alone, had no follow-up or had non-atherosclerotic myocardial injury were excluded.

Outcome and Candidate Predictors: The primary outcome was 12-month MACE, defined as a composite of all-cause mortality, recurrent myocardial infarction, ischemic stroke, target-vessel revascularization or heart failure hospitalization. Predictors that were tested included age, sex, hypertension, diabetes, smoking, previous coronary disease, chronic kidney disease, Killip class, location of the infarct, symptom-to-door time, door-to-balloon time, systolic blood pressure, heart rate, haemoglobin, creatinine, glucose, troponin, lipid variables, LVEF, single versus multivessel disease, culprit artery, thrombus burden, final TIMI flow, stent length, and discharge medications and adherence risk documented at first follow-up.

Model Development: The data was split into training and testing sets (70:30) with stratification according to outcome. 42 baseline variables were used to train logistic regression, random forest and gradient boosting. Winsorization of extreme percentiles was performed for continuous predictors. Median or mode was used for missingness < 6%. Five-fold cross validation was performed in the training cohort to tune hyperparameters.

The performance and subgroup analysis are provided. Performance and subgroup analysis are available.

AUC, sensitivity, specificity, positive predictive value, negative predictive value, Brier score and calibration slope were used to evaluate model performance. Subgroup performance was estimated for age <60 years, sex, diabetes, anterior STEMI, chronic kidney disease and multivessel disease. The baseline comparisons were considered statistically significant when the p value was < 0.05.

Results

Sixty-eight hundred STEMI patients were included in the final cohort. Mean age was 58.9 ± 11.6 years and 74.1% were male. Anterior wall STEMI was seen in 46.6%, diabetes in 34.0%, chronic kidney disease in 12.8% and multivessel disease in 38.5%. Twelve-month MACE was seen in 118 patients (17.4%), including 42 deaths, 31 hospitalizations for heart failure, 18 recurrent myocardial infarctions, 11 strokes and 16 target-vessel revascularizations.

Patients with MACE had higher risk baseline features such as older age, diabetes, renal dysfunction, higher Killip class, lower EF and longer symptom-to-balloon time. Gradient boosting was the best discriminator and had an acceptable calibration. The model was well performing in most of the subgroups, with a slight drop in discrimination in the chronic kidney disease and older age groups (60 years and older). (Table 1-3)

Table 1: Baseline characteristics of STEMI patients according to 12-month MACE

Variable	MACE (n=118)	No MACE (n=562)	p value
Age (years)	64.3 ± 10.8	57.8 ± 11.4	<0.001
Male sex	82 (69.5%)	422 (75.1%)	0.207
Diabetes mellitus	56 (47.5%)	175 (31.1%)	0.001
Chronic kidney disease	28 (23.7%)	59 (10.5%)	<0.001
Anterior STEMI	67 (56.8%)	250 (44.5%)	0.015
Killip class II-IV	49 (41.5%)	110 (19.6%)	<0.001
LVEF (%)	39.8 ± 8.7	46.1 ± 9.4	<0.001

Table 2: Comparative performance of prediction models for 12-month MACE

Model	AUC (95% CI)	Sensitivity	Specificity	Brier score
Logistic regression	0.78 (0.72-0.84)	70.3%	72.1%	0.132
Random forest	0.82 (0.76-0.87)	75.4%	75.8%	0.119
Gradient boosting	0.84 (0.79-0.89)	78.0%	77.4%	0.111

Table 3: Subgroup-specific AUC of the gradient boosting model

Subgroup	n	MACE rate	AUC
Age <60 years	348	12.1%	0.85
Age ≥60 years	332	22.9%	0.80
Male	504	16.3%	0.84
Female	176	20.5%	0.81
Diabetes present	231	24.2%	0.80
No diabetes	449	13.8%	0.86
Chronic kidney disease	87	32.2%	0.77
Multivessel disease	262	24.0%	0.82

Discussion

The results of this study show that machine learning models, which integrate clinical, laboratory, angiographic and discharge treatment variables, can enhance the prediction of 12-month MACE after STEMI. Gradient boosting had the best AUC and the lowest Brier score, indicating that it discriminated and estimated the probability better than the conventional logistic regression in this dataset. The results are consistent with previous studies showing that flexible models are able to model complex interactions with risk in acute coronary syndrome [9,10].

Clinically credible were the most important predictors. Killip class and LVEF are markers of acute hemodynamic compromise and ventricular reserve, whereas creatinine and chronic kidney disease are markers of cardio-renal vulnerability, diabetes and multivessel disease are markers of diffuse atherosclerosis, and symptom-to-balloon time is a marker of ischemic burden. These are well known to be associated with adverse prognosis following myocardial infarction [11,12].

One of the strengths of the present study is the ability to perform a subgroup analysis. Overall discrimination was satisfactory, but poorer in the chronic kidney disease and older patients. Calibration and subgroup reporting are critical since a seemingly good global model may not be as reliable in patients that are commonly seen in clinical practice [13-15]. Failure to recognise such variation may result in inappropriate reassurance or unnecessary intensive follow up.

Prediction models should be used in the support of clinical pathways including early follow-up, heart failure surveillance, cardiac rehabilitation intensity, adherence counselling and optimization of secondary prevention. They should not be used as a substitute for clinical judgment and/or guideline-directed therapy. Claims of publication need to be supported by transparent reporting and external validation before they can be used to support clinical implementation [16].

The limitations are retrospective design, single-centre data set, lack of external validation and potential documentation bias of adherence risk. Sample size was sufficient for model development but larger multicentre samples would provide more robust subgroup calibration.

Conclusion

In this retrospective cohort, a gradient boosting prediction model had superior 12-month MACE discrimination following STEMI compared with logistic regression and random forest. Killip class, LVEF, renal function, diabetes, symptom-to-balloon time and angiographic severity were important predictors. Performance of the subgroups differed, especially for older age and chronic kidney disease, highlighting the importance of calibration, external validation and reporting of subgroups prior to clinical use.

References

1. Ibanez B, James S, Agewall S, Antunes MJ, Bucciarelli-Ducci C, Bueno H, et al. 2017 ESC guidelines for management of acute myocardial infarction in patients presenting with ST-segment elevation. Eur Heart J. 2018;39(2):119-177. doi:10.1093/eurheartj/ehx393. PMID:28886621.
2. O'Gara PT, Kushner FG, Ascheim DD, Casey DE Jr, Chung MK, de Lemos JA, et al. 2013 ACCF/AHA guideline for management of ST-elevation myocardial infarction. Circulation. 2013;127(4):e362-e425. doi:10.1161/CIR.0b013e3182742cf6. PMID:23247304.
3. D'Agostino RB Sr, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care. Circulation. 2008;117(6):743-753. doi:10.1161/CIRCULATIONAHA.107.699579. PMID:18212285.
4. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and validation. Eur Heart J. 2014;35(29):1925-1931. doi:10.1093/eurheartj/ehu207. PMID:24898551.
5. Collins GS, Reitsma JB, Altman DG, Moons KGM. TRIPOD statement for transparent reporting of prediction models. Ann Intern

Med. 2015;162(1):55-63. doi:10.7326/M14-0697. PMID:25560714.
6. Chen W, Tan Y, Wang Y, Wang S, Huang Y, Liu J, et al. Prediction models for major adverse cardiovascular events following ST-segment elevation myocardial infarction and subgroup-specific performance. BMC Cardiovasc Disord. 2023;23:241. doi:10.1186/s12872-023-03250-x. PMID:37189118.
7. Chopannejad S, Granger CB, Mehta RH, Lopes RD, Al-Khatib SM, Pokorney SD. Predicting major adverse cardiovascular events in acute coronary syndrome using machine learning. J Am Heart Assoc. 2022;11:e025997. doi:10.1161/JAHA.122.025997. PMID:35876619.
8. Liu Z, Liang J, Zhang L, Zhang Y, Wang C, Chen Z, et al. Prediction of major adverse cardiovascular events following acute myocardial infarction using cardiac obesity marker and clinical variables. BMC Cardiovasc Disord. 2024;24:724. doi:10.1186/s12872-024-04409-6. PMID:39732819.
9. Fox KAA, Dabbous OH, Goldberg RJ, Pieper KS, Eagle KA, Van de Werf F, et al. Prediction of risk of death and myocardial infarction in the six months after presentation with acute coronary syndrome: GRACE risk score. BMJ. 2006;333(7578):1091. doi:10.1136/bmj.38985.646481.55. PMID:17032691.
10. Antman EM, Cohen M, Bernink PJLM, McCabe CH, Horacek T, Papuchis G, et al. TIMI risk score for unstable angina/non-ST elevation MI. JAMA. 2000;284(7):835-842. doi:10.1001/jama.284.7.835. PMID:10938172.
11. Stone GW, Maehara A, Lansky AJ, de Bruyne B, Cristea E, Mintz GS, et al. A prospective natural-history study of coronary atherosclerosis. N Engl J Med. 2011;364(3):226-235. doi:10.1056/NEJMoa1002358. PMID:21247313.
12. Yeh RW, Sidney S, Chandra M, Sorel M, Selby JV, Go AS. Population trends in incidence and outcomes of acute myocardial infarction. N Engl J Med. 2010;362(23):2155-2165. doi:10.1056/NEJMoa0908610. PMID:20558366.
13. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. J Mach Learn Res. 2011;12:2825-2830. doi:10.5555/1953048.2078195.
14. Austin PC, Steyerberg EW. Events per variable and prediction model performance. Stat Methods Med Res. 2017;26(2):796-808. doi:10.1177/0962280214558972. PMID:25301842.
15. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Calibration: the Achilles heel of predictive analytics. BMC Med. 2019;17:230. doi:10.1186/s12916-019-1466-7. PMID:31842878.
16. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis. Ann Intern Med. 2015;162(1):W1-W73. doi:10.7326/M14-0698. PMID:25560730.