BMC Medical Informatics and Decision Making

RESEARCH

Open Access

Risk factor analysis and predictive nomogram development for in-hospital mortality in patients with ST-segment elevation myocardial infarction

Tahereh Roostami[1], Maryam Farhadian[1,2*], Amirhossein Yazdi[3] and Hossein Mahjub[1,2]

Abstract

Background Identifying predictors of in-hospital mortality in patients with ST-segment elevation myocardial infarction (STEMI) is a major concern in cardiology. The aim of this study was to identify risk factors and develop a nomogram to predict in-hospital mortality in STEMI patients.

Materials and methods This single-center study was a retrospective cohort analysis of all STEMI patients consecutively referred to Farshchian Hospital (Hamadan Province-Iran) from April 2021 to August 2024. Four different feature selection methods were used to select common important variables in the prediction model: Boruta, Recursive Feature Elimination (RFE), Random Forest (RF) and LASSO. The uneven distribution of the different classes 2,356 (91.7%) alive and 212 (8.3%) dead) was dealt with using the SMOTE method. After splitting the data into a training (70%) and a test (30%) dataset, a multiple logistic regression model was formulated using the significant variables identified. A nomogram predicting in-hospital mortality was then constructed and validated.

Results The findings indicate that age (OR = 1.05: 95% CI: 1.03–1.06), gender (female OR = 1.71: 95% CI: 1.22–2.41), length of stay (OR = 0.89: 95% CI: 0.83–0.96), blood urea nitrogen level (OR = 1.02: 95% CI: 1.00-0.03), white blood cell count (OR = 1.07: 95% CI: 1.03–1.1), Creatinine (OR = 1.74: 95% CI: 1.36–2.22), fasting blood glucose (OR = 1.007: 95% CI: 1.005–1.009), uric acid (OR = 1.07: 95% CI: 1.02–1.12), potassium (OR = 1.2: 95% CI: 0.95–1.52) and systolic blood pressure (OR = 0.98: 95% CI: 0.97–0.99) are pivotal factor in predicting in-hospital mortality in STEMI patients. The predictive model demonstrated high accuracy (84%) and excellent discriminatory ability with an AUC of 0.91. The calibration plot demonstrated the model's strong discriminatory performance in distinguishing between the two classes.

Conclusions The development of a nomogram for reliably predicting in-hospital mortality in STEMI patients provides clinicians with a practical visual aid for identifying high-risk patients. By enabling tailored care strategies, this tool improves therapeutic precision and ultimately leads to better clinical outcomes.

Clinical trial number Not applicable.

*Correspondence:
Maryam Farhadian
maryam_farhadian80@yahoo.com

Full list of author information is available at the end of the article

Keywords Acute myocardial infarction, Nomogram, In-hospital mortality

Background

Cardiovascular disease (CVD) is responsible for approximately 30% of all deaths. Among the various cardiovascular diseases, acute myocardial infarction (AMI) remains the leading cause of death and hospitalization on a global scale. Despite advances in cardiac interventions, AMI remains one of the most common life-threatening diagnoses in hospital emergency departments, leading to increased hospitalizations costs, long-term disability and mortality [1, 2].

The age-standardized incidence rate of AMI in Iran has been reported to be 73. 3 per 100,000 populations. This figure indicates a significant burden of AMI in the Iranian population. The prevalence of ST-segment elevation myocardial infarction (STEMI) in Iran is almost higher than that of NSTEMI [3]. While both STEMI and NSTEMI contribute to CVD mortality, STEMI is associated with significantly higher in-hospital mortality rates in Iran, necessitating tailored risk-assessment tools for this population [4, 5]. Thus, our study focuses exclusively on STEMI patients to develop a nomogram to predict in-hospital mortality among this patients.

In-hospital mortality refers to deaths that occur during hospitalizations and varies with disease and quality of care. In-hospital mortality is a major concern in cardiology for patients with STEMI, and several studies have investigated different prediction models. Mortality rates can vary significantly according to geographical location and patient demographics, requiring appropriate approaches to risk assessment. In-hospital mortality rates also vary according to a number of factors, including gender, time to treatment and clinical characteristics [6, 7].

The in-hospital mortality rate for STEMI patients in different cohorts in Iran has been reported to be approximately 6–14%, depending on the population and treatment protocols. Recent data suggest that in-hospital mortality rates have been decreasing over time [4, 8]. The fact that mortality rates from STEMI vary greatly depending on the setting, as well as being influenced by factors such as reperfusion timing, comorbidities and healthcare resources, highlights the need for robust predictive tools as a critical aspect of cardiovascular care. Various scoring algorithms and machine learning models have been developed to improve the accuracy of these predictions [6, 7, 9].

Statistical models to predict the risk of disease or other adverse events are being developed in many clinical areas. These models are intended to help patients and clinicians make informed decisions. Nomograms are simple graphical representations that aim to convey knowledge in an understandable way, enabling clinicians to derive predicted response values for a particular patient by manually connecting points on the corresponding nomogram. To develop a robust nomogram, however, the identification of effective risk factors is essential [10–14].

The identification of risk factors for in-hospital mortality in STEMI patients is imperative for enhancing patient outcomes and guiding treatment strategies. Extant research has identified several pivotal factors contributing to in-hospital mortality, which can inform clinical practices and enhance risk stratification. The effective identification of risk factors is, thus, vital not only for the timely making of clinical decisions, but also for the subsequent development of long-term treatment protocols and the enhancement of patient outcomes.

The present study therefore sought to identify the most significant factors affecting in-hospital mortality in STEMI patients and develop a predictive nomogram combined with feature selection algorithms.

Materials and methods

Study design and population

The present study constitutes a retrospective cohort analysis of a single center. It is an investigation of all STEMI patients who were consecutively referred to Farshchian Hospital in Iran from April 2021 to August 2024.

Patients admitted to hospital with ICD-10 codes I21.0, I21.1, I21.2 and I21.9 were identified. Exclusion criteria were death before hospital admission and incomplete patient records.

The study population encompassed patients who had attained over 18 years of age and who had been admitted to hospital with a diagnosis of STEMI. Definition of STEMI does adhere to internationally accepted clinical guidelines. According to the most recent standards from the American Heart Association (AHA), American College of Cardiology (ACC), and European Society of Cardiology (ESC), the electrocardiographic criteria for STEMI require ST-segment elevation at the J point in two contiguous leads. Specifically, for all leads except V2-V3, the threshold should be ≥ 0.1 mV (1 mm), while for V2–V3, the sex/age-adjusted thresholds: men ≥ 40 years should have ≥ 0.2 mV (2 mm), men < 40 years ≥ 0.25 mV (2.5 mm), and women ≥ 0.15 mV (1.5 mm) [15].

All stages were developed according to the conditions set by the Iranian Hospital Ethics Committee and ethical standards were followed. The names of the subjects were kept confidential. Ethical approval code: IR.UMSHA.REC.1402.349 was obtained from the Ethics Committee of Hamadan University of Medical Sciences to obtain information from cardiac patient records and use this information for Hamadan Farshchian Hospital.

The potential predictors of mortality were retrieved from the database, as per previous studies. This included demographic information, past medical history, initial symptoms and laboratory findings based on clinical characteristics on the first day of admission. The primary outcome of interest was in-hospital mortality. A flowchart summarizing the study process was presented in Fig. 1.

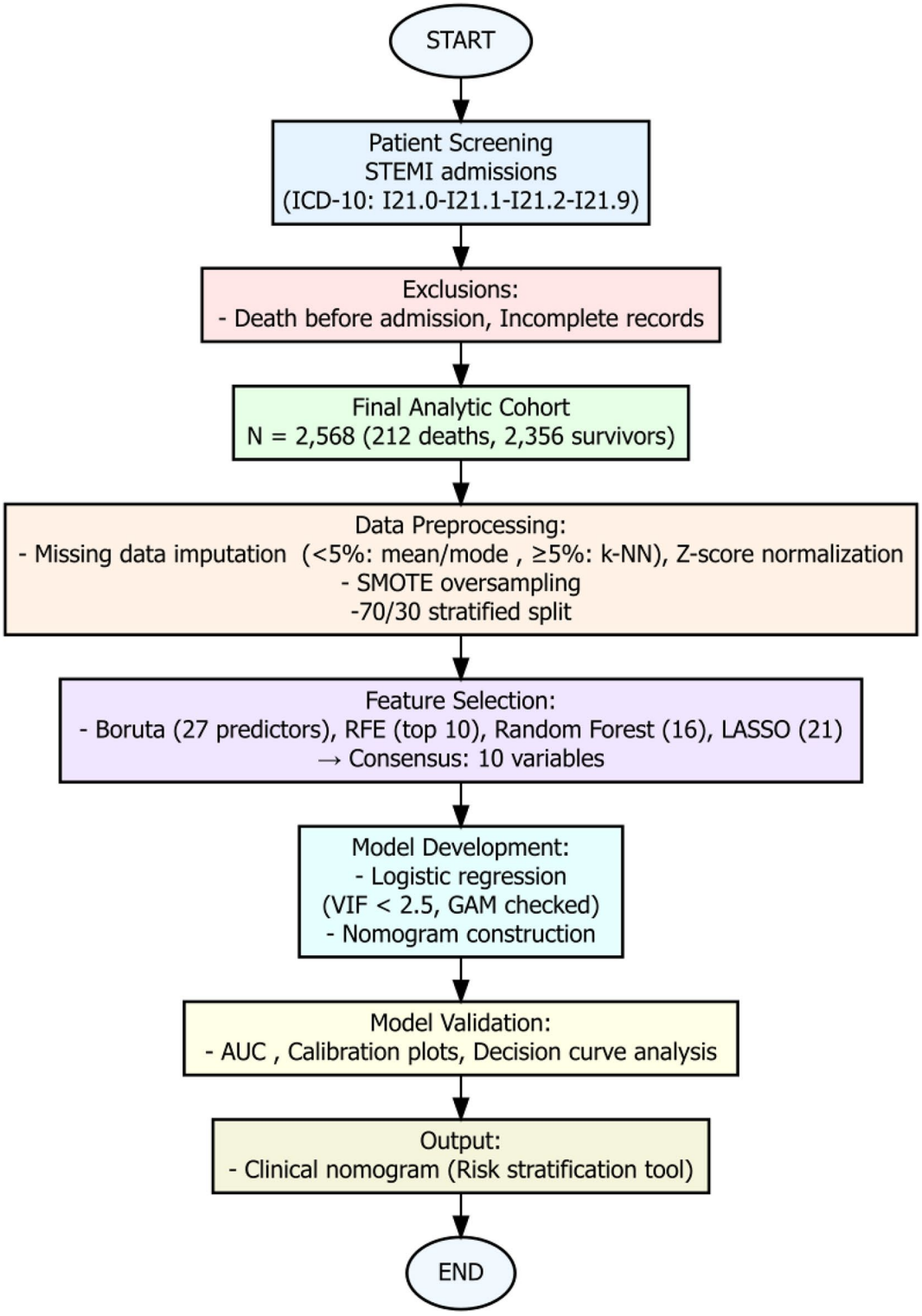

Fig. 1 Flow chart of the study design

Data preprocessing

In the first pre-processing step, after removing irrelevant values, missing values for variables with very low missing percentages (less than 5%) were filled with the mean or most frequent value corresponding to the variable. For other variables with missing percentages higher than 5%, the k-nearest neighbor's method was used. All continuous variables were standardized using Z-score normalization prior to feature selection. Categorical variables (e.g., sex, comorbidities) were kept as binary (0/1) values.

Imbalanced data

The distribution of the classes (occurrence or non-occurrence of hospital death in patients) shows that the number of samples in class 0 or non-occurrence of death (2356) is significantly higher than in class 1 or occurrence of death (212). This imbalance can lead to problems in the prediction accuracy of the model, as the model may have a tendency to predict the majority class more. The SMOTE (Synthetic Minority Oversampling Technique) technique was used to deal with the class imbalance.

SMOTE is a widely utilized method for addressing class imbalance in datasets, particularly in classification problems. The SMOTE aims to balance class distributions by generating synthetic examples of the minority class rather than simply duplicating existing instances. Initially, the k-nearest neighbors of the minority class samples are determined for each sample of the minority class. Then, for each minority class sample, one of the neighbors is randomly selected. Utilizing interpolation, a new sample is created between the two aforementioned samples. Following this process, the data is reconstructed [16, 17].

SMOTE implementation included quality control measures assessing synthetic sample positioning within original feature space. We evaluated potential noise introduction through nearest neighbor analysis and multidimensional outlier detection. Over fitting risks were mitigated through separate validation sets.

Feature selection

The identification of risk factors is an essential step in the design of a predictive nomogram of in-hospital mortality in STEMI patients. Boruta feature selection techniques, Recursive Feature Elimination, LASSO Regularization and Random Forest were used to identify the most important predictor of in-hospital death in STEMI patients.

We selected these four methods to present a wide range of methodological approaches including filtering techniques (Boruta), wrapper technique (RFE), and embedded techniques (LASSO and random forest).

Boruta

The Boruta method is a feature selection technique that enhances classification algorithms, particularly Random Forest, through the evaluation of feature significance relative to "shadow" features, which are artificially generated. For each feature present in the original dataset, Boruta generates a corresponding "shadow" feature through a randomization of its values, thereby ensuring that the "shadows" bear no correlation to the target variable. The algorithm subsequently trains classifiers on this amalgamation of the original features and their corresponding "shadows", and calculates importance scores for both feature types.

In each iteration, Boruta compares the importance of the original features to that of the most significant shadow feature, with features that demonstrate significantly greater importance than the best shadow feature being classified as "Confirmed." In turn, those with a lesser importance are marked as "Rejected." Those that do not clearly fit into either of these categories are labeled as "Tentative". This classification process continues until all features are sorted into one of the three categories or until a specified maximum number of iterations are reached [18].

Recursive feature elimination

Recursive Feature Elimination (RFE) is a technique used to select features in a data set by gradually eliminating the least significant ones to improve model performance. The RFE process starts by training a model using all available features. The model then calculates an importance score for each feature, reflecting its importance in making predictions. These scores may be obtained from a variety of sources, including the coefficients employed within linear models or the feature importance metrics utilized in tree-based models.

Once the importance scores have been determined, the RFE will rank the features and identify the least important features to be removed. The model is then retrained without these features and new importance scores are calculated. This iterative process continues, removing one or more features at each step, until either a specified number of features remain or further removals negatively affect model performance. The decision to retain or discard features depends on their impact on the performance metrics; if removing a feature causes a significant drop in performance, it is retained; otherwise, it is discarded [19].

LASSO regularization

Least Absolute Shrinkage and Selection Operator (LASSO) regularization is a powerful technique used in machine learning for both regularization and feature selection, particularly in classification tasks. LASSO can

effectively reduce the number of features used in a model, improving interpretability and reducing the risk of overfitting, by adding an L1 penalty to the loss function [20].

Random forest

Random forest is an ensemble learning technique that is excellent at the evaluation of the importance of features. It consists of multiple decision trees, where each tree is made up of internal nodes and leaves. At each internal node, a selected feature determines how to split the data set into two groups with similar results. The features for these splits are selected based on specific criteria, such as Gini impurity or information gain for classification tasks.

Feature importance is assessed using Gini importance, which measures the average reduction in impurity that a feature contributes when used to split nodes across all trees in the random forest. The more a feature helps to reduce impurity - whether through Gini impurity or entropy - the more important it becomes in the model. Thus, features that lead to a significant reduction in impurity are considered more important [21].

Logistic regression

Following the identification of important predictor variables in the feature selection stage, a multiple logistic regression model was fitted to create a nomogram. Multicollinearity was assessed through variance inflation factors (VIF), with all values < 2.5 indicating acceptable levels. Pairwise correlations among predictors were examined, with the strongest correlation ($r = 0.4$) between Cr and BUN (Fig. 2). For sensitivity analysis three models were built: one with both Creatinine and BUN; one with only Creatinine or BUN; and one with the BUN/Creatinine ratio as an alternative. After comparing their

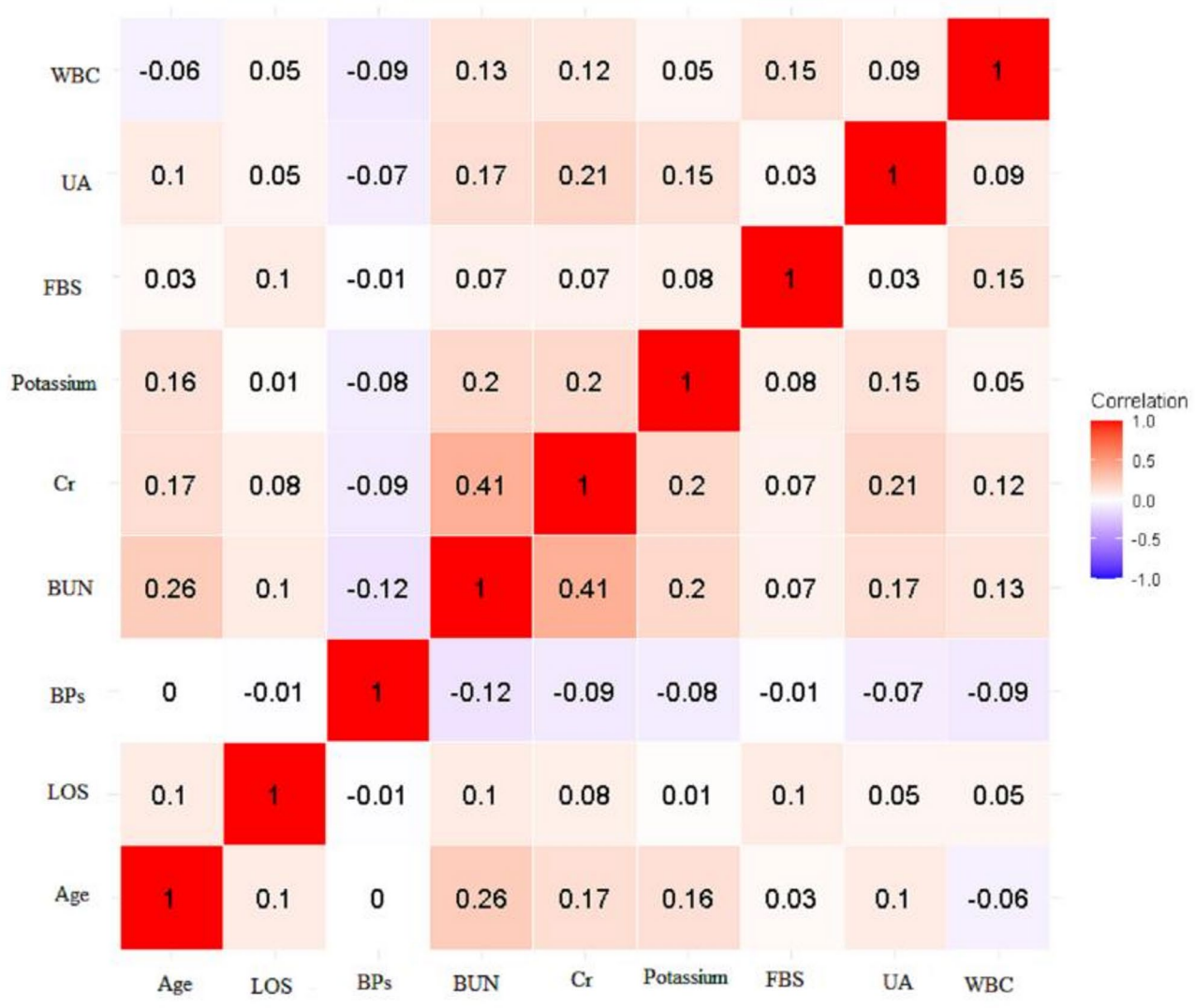

Fig. 2 Pairwise correlation matrices computed for continuous variables

AIC/BIC, C-statistic (discrimination) and calibration, we found that the different models did not differ significantly from each other. Also, effect sizes and significance levels remained consistent. Although the performance of the third model was slightly better, this new variable was not significant in multiple logistic model. Therefore, we decided to create a nomogram incorporating both variables.

Also, the linearity of continuous variables was assessed using the Generalised Additive Models (GAM) method and deviance residual analysis. The results confirmed a linear relationship between the outcome and continuous variables.

Nomogram construction and evaluation

Following the application of feature selection techniques to determine key risk factors linked to in-hospital mortality among STEMI patients, a predictive nomogram using these critical variables was developed. Model performance was assessed using metrics such as area under the curve (AUC) and calibration plots using 30% of the data as test data sets, and clinical utility was assessed using decision curve analysis (DCA).

The 70:30 stratified split was chosen to balance model development needs with validation rigor, supported by sensitivity analyses comparing alternative approaches. While this provides robust internal validation, we are currently suggesting external validation across multiple Iranian centers to assess generalizability in future study.

We conducted our analysis in R statistical software using the following packages: "Boruta" and "caret" (for feature selection), "randomForest" (for a random forest implementation), "smotefamily" for class imbalance correction, "glmnet" for regularised regression, "rms" for advanced regression strategies and "rmda" for decision curve analysis.

Results

Our study analyzed 2,568 consecutive STEMI patients, of whom 212 (8.3%) died during hospitalization, while 2,356 (91.7%) survived to discharge. Table 1 details baseline characteristics stratified by survival status.

Gender demonstrated a statistically significant association with mortality risk. Among 633 female patients, survival rates were 85.9% ($n = 544$), with 14.1% mortality ($n = 89$). In contrast, male patients showed higher survival rates at 93.6% ($n = 1,812$), with only 6.4% mortality ($n = 123$).

As shown in Table 2 age differences between groups were pronounced: deceased patients averaged 64.16 ± 14.01 years, significantly older than survivors (41.16 ± 44.26 years). No significant correlation was observed between hospitalization duration (LOS) and mortality risk.

Univariate analysis revealed significant mortality associations with: pre-existing heart disease history, family history of heart disease, diabetes history and smoking history. Hemodynamic differences were also noted: systolic blood pressure averaged 121.90 ± 30.68 mmHg in deceased patients versus 138.27 ± 26.24 mmHg in survivors.

No mortality associations were identified for: activity-related pain exacerbation, Dyspnea, Nausea/vomiting, Headache, Pulmonary disease history, PTT levels, Triglyceride/HDL profiles, Hemoglobin counts, Platelet counts (PLT) and mean corpuscular hemoglobin (MCH).

Firstly the Boruta algorithm was applied to our dataset to determine statistically significant features, using Random Forest as its core classification method. This analysis revealed 27 key features that outperformed their randomly generated "shadow" counterparts, demonstrating clear predictive importance. Among these were clinically relevant predictors of mortality, including length of stay (LOS), creatinine (Cr) levels, white blood cell count (WBC), uric acid (UA), systolic blood pressure (BPs), potassium, fasting blood sugar (FBS), blood sugar (BS), age, and others. Conversely, 20 features were deemed non-essential for prediction, while three attributes (chronic kidney disease status [CKD], LDL cholesterol, and total cholesterol [TC]) showed borderline significance. These findings are visualized in Fig. 3.

Following this, RFE was used to identify the most relevant features. The analysis revealed that retaining 10 features produced the highest accuracy score (0.938) when validated through 10-fold cross-validation. These selected features include LOS, Cr, WBC, UA, blood urea nitrogen (BUN), BPs, FBS, Potassium, Age, and BS (Fig. 4).

The random forest model identified sixteen key predictors in descending order of importance: LOS, Cr, WBC, UA, BUN, FBS, Potassium, BPs, Age, PLT, BS, PR, MCV, BPd, MCHC, and O2sat. The variable importance ranking is visualized in the accompanying Fig. 5.

To determine the most impactful predictors, we employed also the LASSO method. This approach identified 21 key features strongly associated with in-hospital mortality. Notably, serum Cr emerged as the most influential variable, with the highest coefficient value of 0.425. The LASSO coefficient profiles and cross-validation results for tuning parameter optimization are visualized in Fig. 6, which demonstrates the model's selection process and validation strategy through ten-fold cross-validation.

Ten common variables emerged as key predictors of mortality: age, gender, LOS, BUN, WBC, Cr, FBS, UA, potassium, and BPs (Fig. 7).

The logistic regression analysis (illustrated in the Fig. 8) revealed odds ratios and 95% confidence intervals for

Table 1 Baseline characteristics of the population cohort of STEMI patients

		Total		Alive		Death		*P*-value*
		N	%	*N*	%	*N*	%	
Sex	Female	633	24.6	544	23.1	89	42.0	< 0.001
	Male	1935	75.4	1812	76.9	123	58.0	
ICD-10 Diagnosis	I21.9	204	7.9	161	6.8	43	20.3	< 0.001
	I21.1	1051	40.9	971	41.2	80	37.7	
	I21.2	42	1.6	40	1.7	2	0.9	
	I21.0	1271	49.5	1184	50.3	87	41.0	
Chest Pain	No	84	3.3	70	3.0	14	6.6	0.004
	Yes	2484	96.7	2286	97.0	198	93.4	
SOB (Shortness of Breath)	No	1531	59.6	1416	60.1	115	54.2	0.096
	Yes	1037	40.4	940	39.9	97	45.8	
Edema	No	2515	97.9	2312	98.1	203	95.8	0.020
	Yes	53	2.1	44	1.9	9	4.2	
Headache	No	2341	91.2	2144	91.0	197	92.9	0.345
	Yes	227	8.8	212	9.0	15	7.1	
Cold Sweat	Yes	1321	51.4	1232	52.3	89	42.0	< 0.001
	No	490	19.1	457	19.4	33	15.6	
	Non	757	29.5	667	28.3	90	42.5	
HTN (Hypertension)	Positive	1142	44.5	1027	43.6	115	54.2	0.002
	Negative	1404	54.7	1311	55.6	93	43.9	
	Non	22	0.9	18	0.8	4	1.9	
Diabetes	Positive	615	23.9	534	22.7	81	38.2	< 0.001
	Negative	1925	75.0	1800	76.4	125	59.0	
	Non	28	1.1	22	0.9	6	2.8	
History of Heart Disease	Positive	559	21.8	506	21.5	53	25.0	0.095
	Negative	1897	73.9	1752	74.4	145	68.4	
	Non	112	4.4	98	4.2	14	6.6	
Family History of Heart Disease	Positive	419	16.3	406	17.2	13	6.1	< 0.001
	Negative	1181	46.0	1066	45.2	115	54.2	
	Non	968	37.7	884	37.5	84	39.6	
Smoking	Positive	1149	44.7	1081	45.9	68	32.1	< 0.001
	Negative	1115	43.4	1005	42.7	110	51.9	
	Non	304	11.8	270	11.5	34	16.0	
CKD (Chronic Kidney Disease)	Positive	80	3.1	68	2.9	12	5.7	0.026
	Negative	2488	96.9	2288	97.1	200	94.3	
COPD (Chronic Obstructive Pulmonary Disease(	Positive	76	3.0	66	2.8	10	4.7	0.115
	Negative	2492	97.0	2290	97.2	202	95.3	
Troponins	Positive	1387	54.0	1225	52.0	162	76.4	< 0.001
	Negative	1181	46.0	1131	48.0	50	23.6	
PCI	No	667	26.0	528	22.4	139	65.6	< 0.001
	Yes	1901	74.0	1828	77.6	73	34.4	
CABG	No	2505	97.5	2308	98.0	197	92.9	< 0.001
	Yes	63	2.5	48	2.0	15	7.1	
Pericardiotomy	No	2508	97.7	2310	98.0	198	93.4	< 0.001
	Yes	60	2.3	46	2.0	14	6.6	
Cuttrezasion	No	2522	98.2	2321	98.5	201	94.8	< 0.001
	Yes	46	1.8	35	1.5	11	5.2	
CPR	No	2346	91.4	2326	98.7	20	9.4	< 0.001
	Yes	222	8.6	30	1.3	192	90.6	

Table 2 Clinical characteristics of STEMI patients on the first day of admission to hospital

	Alive	Death	*P*-value*
	Mean ± SD	Mean ± SD	
Age (year)	62.38 ± 12.09	71.18 ± 12.23	< 0.001
Length of Stay (LOS)	3.59 ± 2.31	3.46 ± 3.82	0.445
Systolic Blood Pressure (BPs)	138.27 ± 26.24	121.90 ± 30.68	< 0.001
Diastolic Blood Pressure (BPd)	87.64 ± 23.01	78.44 ± 21.42	< 0.001
Pulse Rate (PR)	80.24 ± 16.27	87.41 ± 24.07	< 0.001
Heart Rate (HR)	79.99 ± 16.57	87.36 ± 23.80	< 0.001
O2sat	93.57 ± 5.20	90.32 ± 0.89	< 0.001
Prothrombin Time (PT)	13.63 ± 3.49	15.61 ± 10.18	< 0.001
International Normalized Ratio (INR)	1.11 ± 0.28	1.28 ± 0.55	< 0.001
Partial Thromboplastin Time (PTT)	31.45 ± 13.86	32.23 ± 13.72	0.434
Blood Sugar (BS)	160.50 ± 85.09	223.76 ± 134.27	< 0.001
Blood Urea Nitrogen (BUN)	18.24 ± 9.84	27.41 ± 15.20	< 0.001
Creatinine (Cr)	1.06 ± 0.47	1.57 ± 0.88	< 0.001
Na	139.05 ± 5.76	138.13 ± 5.22	0.025
Potassium	4.15 ± 0.51	4.45 ± 0.79	< 0.001
CKMB	90.49 ± 107.06	121.50 ± 120.10	< 0.001
Fasting Blood Sugar (FBS)	137.59 ± 59.57	182.62 ± 97.98	< 0.001
Uric Acid (UA)	5.88 ± 2.34	7.32 ± 4.24	< 0.001
Total Cholesterol (TC)	164.00 ± 57.15	152.72 ± 50.63	0.005
TGs	147.13 ± 90.21	140.35 ± 78.80	0.290
HDL	40.65 ± 12.66	40.23 ± 10.55	0.638
LDL	94.18 ± 43.48	84.94 ± 43.76	0.003
White Blood Cell Count (WBC)	10.12 ± 4.63	13.52 ± 5.08	< 0.001
Red Blood Cell Count (RBC)	4.92 ± 1.13	4.71 ± 1.05	0.010
HGB	14.64 ± 2.39	14.53 ± 7.59	0.630
HCT	42.58 ± 5.03	41.27 ± 6.45	< 0.001
RDW	13.48 ± 1.38	14.04 ± 1.79	< 0.001
PLT	216.28 ± 67.03	218.58 ± 92.50	0.645
MCV	86.99 ± 6.95	88.27 ± 9.24	0.013
MCH	29.86 ± 2.69	29.52 ± 2.79	0.080
MCHC	34.21 ± 2.71	33.22 ± 2.03	< 0.001
Ejection Fraction (EF)	39.34 ± 9.20	33.42 ± 10.79	< 0.001

these predictors. Notably: Elevated Cr and potassium levels were strongly linked to increase in-hospital mortality risk. Women faced higher mortality odds than men. Longer hospitalization correlated with better survival outcomes.

A nomogram was developed to estimate in-hospital death risk using the identified risk factors. Each variable (e.g., age, LOS, BUN) was assigned points, and the total score provided a personalized mortality prediction (Fig. 9).

Decision curve analysis further substantiated the model's clinical relevance, evincing a substantial net benefit in terms of guiding patient care (Fig. 10). The calibration curve demonstrated a strong alignment between the predicted and observed outcomes (Fig. 11).

Discussion

It is imperative to comprehend the factors that influence in-hospital mortality in STEMI patients in order to enhance patient outcomes. The ongoing research into these predictors will facilitate the development of more effective management protocols and, consequently, result in a reduction in mortality rates within this high-risk population. The findings of the present study suggest significant associations between various variables and the occurrence of death in patients with STEMI. These associations can serve as a foundation for a more profound comprehension of the factors influencing patient survival and for the refinement of treatment methodologies.

The findings indicate that factors such as age, sex, LOS, BUN, WBC, Cr, FBS, UA, potassium and BPs are among the most important variables associated with the occurrence of in hospital death in STEMI patients.

In this study, 91.7% of STEMI patients survived, while 8.3% died. This rate varies according to region and

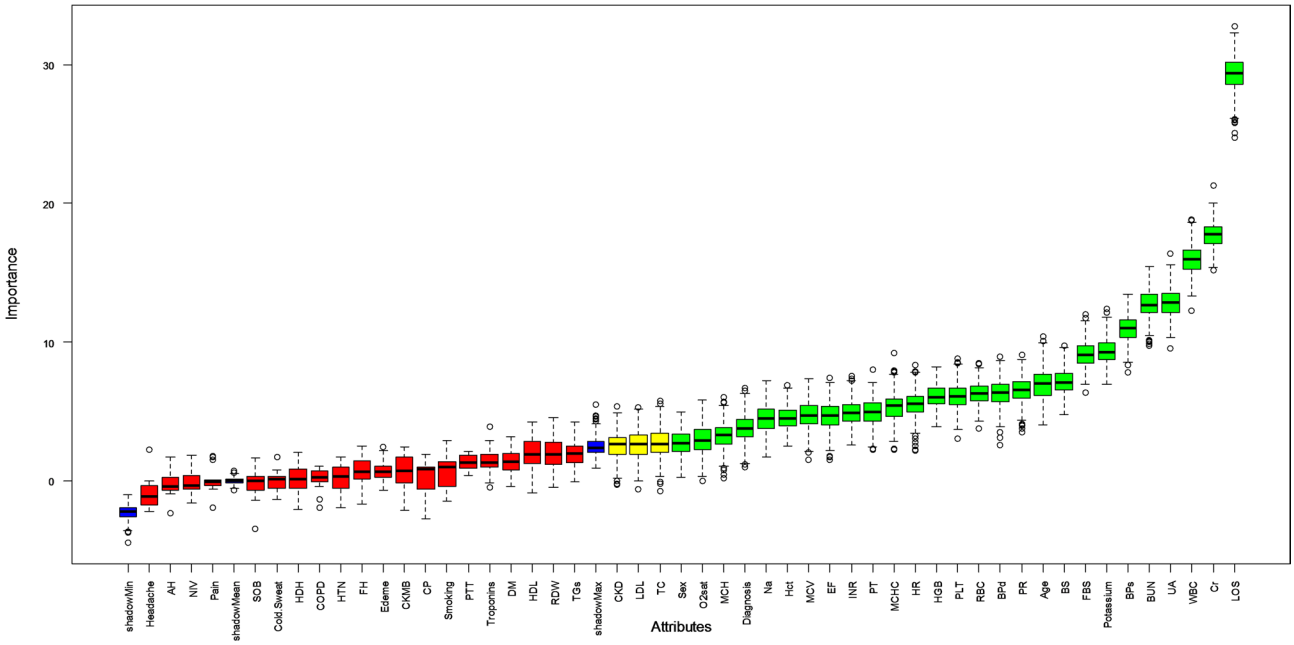

Fig. 3 Importance of each variable according to the Boruta feature selection algorithm

patient management strategies. For example, Momeni et al. showed in a cross-sectional study of the records of 227 patients with AMI that the hospital mortality rate was 8.4% [22]. In a retrospective study of 607 patients with a diagnosis of AMI, followed over a period of up to five years, Farkhani et al. found that 33.6% of patients died during the follow-up period (five years) and 66.4% survived [23].

The results of univariate logistic regression demonstrate that female patients are 2.41 times more likely to die than male patients. Furthermore, studies show that the overall in-hospital mortality rate for women is approximately 17.2%, compared to 9.1% for men (OR = 2.07, 95%CI: 1.85/2.33).The presence of gender differences in mortality is particularly evident in younger patients. For instance, the in-hospital mortality rate for women under 60 years of age was 9.4%, which is almost double the rate for men (4.5%). For women aged 60–69, the mortality rate was 15.2%, compared with 7.5% for men of the same age [24]. Furthermore, women have a higher risk of dying from STEMI than men. This discrepancy may be attributed to a number of factors, including the tendency for female patients to be older and to have more comorbidity at the time of their STEMI, which can complicate treatment and recovery. The observed disparity underscores the necessity for targeted interventions and the development of enhanced treatment strategies for women experiencing STEMI [25].

A significant relationship has been identified between age and the occurrence of death, with the mean age of deceased patients being 71.18 ± 12.23 years, and the mean age of surviving patients being 62.38 ± 12.09 years. This finding indicates an increased risk of death with increasing age in STEMI patients, which could be due to the decline in physiological performance and the increased incidence of comorbidities in the elderly. A number of studies have indicated that patients aged 70 years and over present clinically significantly in cases of AMI and act as an independent predictor of increased risk of in-hospital death [26]. The risk of in-hospital mortality increases with age in STEMI patients. For instance, a study revealed that the mortality rate for patients aged 75 years and over (29.4%) was considerably higher than the overall mortality rate for these patients (17%) [27]. Furthermore, a separate report indicated that the mortality rate for patients aged 80 years and over was 33.6%, while the mortality rate for those younger than 75 years was 11.2% [28].

The present study examined the mean length of hospital stay for patients who died (3.46 ± 3.82 days) and those who survived (3.59 ± 2.31 days).The findings indicated that the length of hospital stay (LOS) for patients with myocardial infarction has a significant impact on long-term outcomes, including readmission rates, mortality, and overall recovery. A prolonged hospitalizations period (4–5 days) facilitates enhanced monitoring and management of complications, a particularly salient benefit for older patients or those with more complex conditions [29, 30].

The mean BPs was significantly lower in deceased patients than in survivors. In addition, studies have indicated that an admission SBP of < 105 mmHg is associated

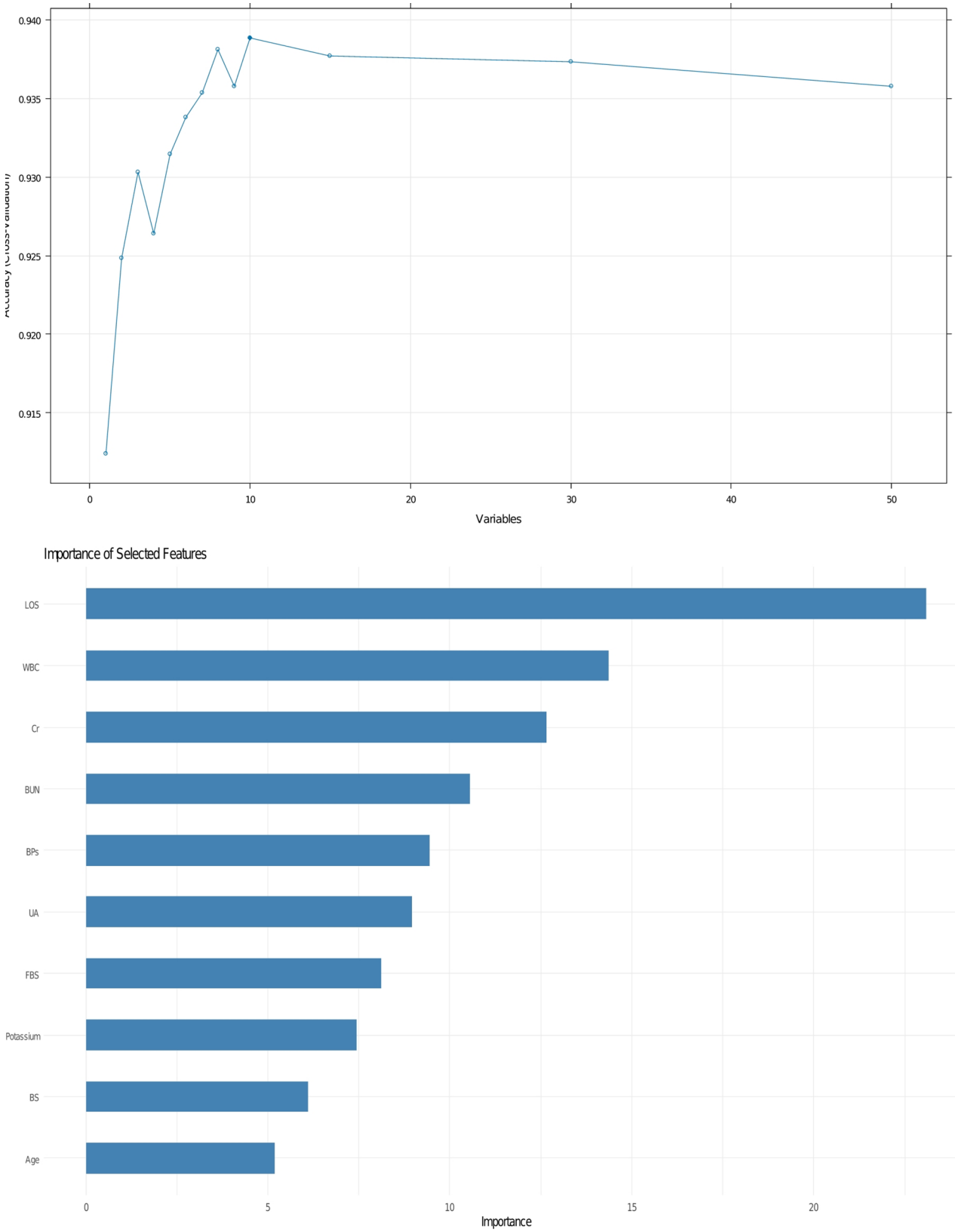

Fig. 4 Recursive Feature Elimination (RFE) identifies 10 features that highly accurately discriminate two classes of STEMI patients (dead-surviving)

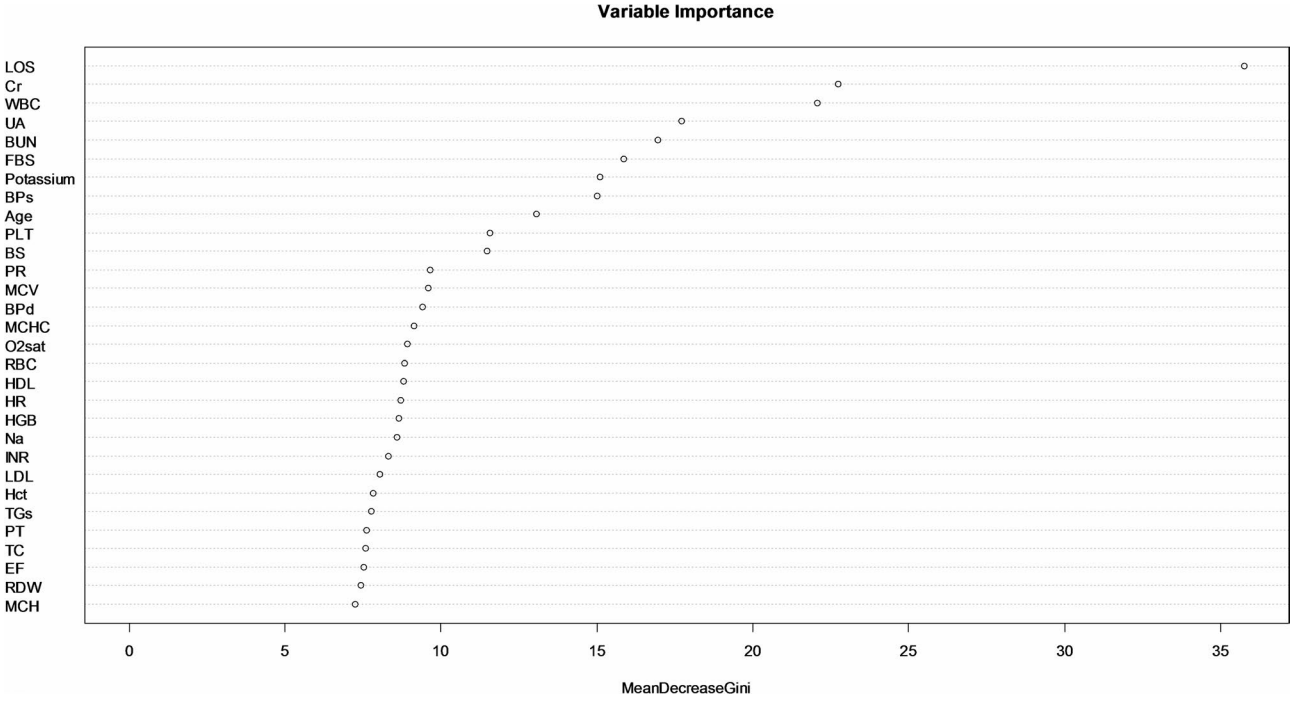

Fig. 5 Random forest features importance

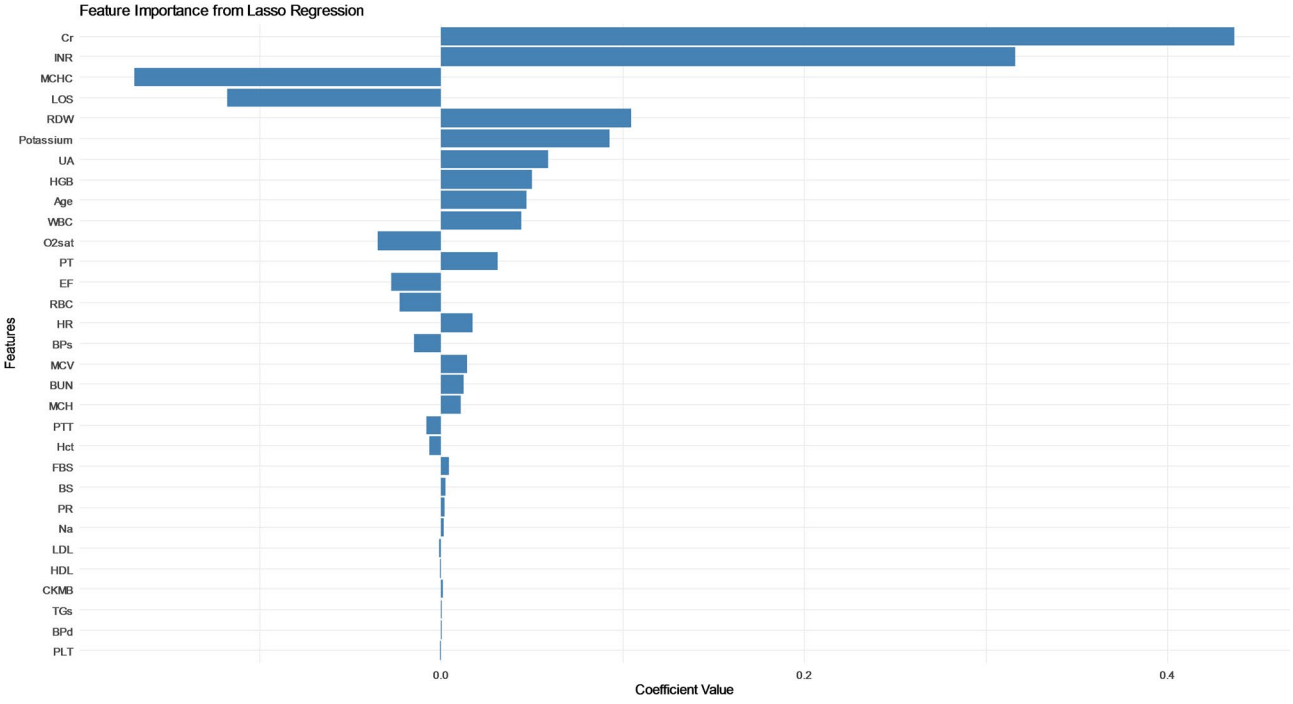

Fig. 6 LASSO regularization feature parameter estimation coefficients

with a higher risk of in-hospital death for STEMI patients. Specifically, it has been shown that patients presenting within the first 48 h after AMI with a mean BPs of less than 125 mmHg are at significantly increased risk of mortality, both during hospitalizations and at one year after the event. The findings emphasize the importance of early identification and management of low BPs in STEMI patients. Interventions aimed at stabilizing blood pressure may improve outcomes and reduce mortality in this vulnerable group [31, 32].

Findings suggest that fasting blood glucose levels may have a significant impact on mortality rates. Higher

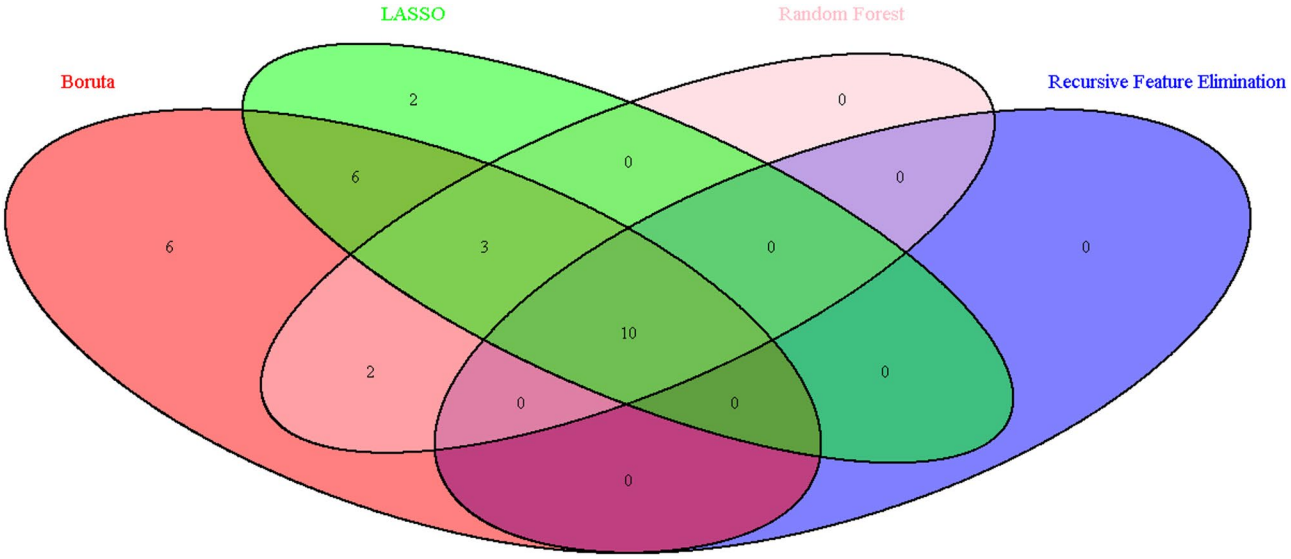

Fig. 7 Venn diagram indicating the overlap of the four feature selection methods

average levels of glucose were observed in patients who died, indicating a potential link between blood glucose levels and mortality. Furthermore, studies have demonstrated a notable association between diabetes mellitus and the mortality rate of STEMI patients. Research consistently indicates that individuals suffering from diabetes have a higher risk of both in-hospital and long-term mortality compared to those without diabetes [33].

The present study's findings indicate that deceased patients exhibited a notably elevated WBC count, which is consistent with previous research. This is particularly interesting given that a negative outcome in acute coronary syndromes can be predicted by an elevated WBC count. Initial leukocyte levels can independently forecast hospital mortality and the onset of heart failure. Therefore, tracking WBC levels during hospitalization may offer critical insights for risk assessment and management strategies for STEMI patients [34–36].

Elevated serum UA levels have been identified as a significant predictor of mortality. Numerous studies indicate that elevated UA levels correlate with adverse outcomes, including increased in-hospital mortality and heart failure, especially in those undergoing PPCI. Furthermore, elevated UA levels have been associated with impaired coronary blood flow during AMI, contributing to elevated mortality rates among patients experiencing concurrently hyperuricemia and inadequate coronary flow during percutaneous PCI [37].

Firstly, it is important to note that this is a single-center, retrospective cohort study. In order to confirm these findings, further research in the form of prospective studies and large multi-center studies is required. Secondly, the development of a nomogram for predicting the likelihood of mortality in STEMI patients demonstrated acceptable predictive capabilities. However, a direct comparison with established risk prediction models, such as the GRACE and TIMI scores, remains challenging due to the absence of relevant data in the patient records, specifically the Killip class and body weight. Moreover, the development of the nomogram in this study was conducted using a particular dataset without undergoing external validation across diverse populations. This restricts the generalizability and applicability of the model in different clinical settings. While our analyses showed acceptable multicollinearity levels, residual confounding between correlated clinical factors may persist.

It is also important to note how potential biases (e.g., selection bias, unmeasured confounders, etc.) may affect the findings. For example, selection bias in this single-center retrospective design favored urban/early-presenting patients, potentially increasing the apparent accuracy of the model. Furthermore, unmeasured confounders and missing key factors, such as pre-hospital delay times, may overestimate the importance of measured clinical factors.

Despite using SMOTE to address class imbalance (8.3% mortality rate), synthetic sample generation may introduce noise. While our validation showed robustness, real-world performance in extreme minority subgroups (e.g., very young patients) requires further testing. Although four methods (Boruta, RFE, RF, LASSO) identified overlapping predictors (e.g., Cr, age), discrepancies in variable rankings (e.g., potassium's significance in LASSO but not RF) highlight the challenge of defining universally "essential" features. While we employed k-nearest neighbors for variables with >5% missingness (e.g., LDL cholesterol), residual bias from incomplete records cannot be ruled out.

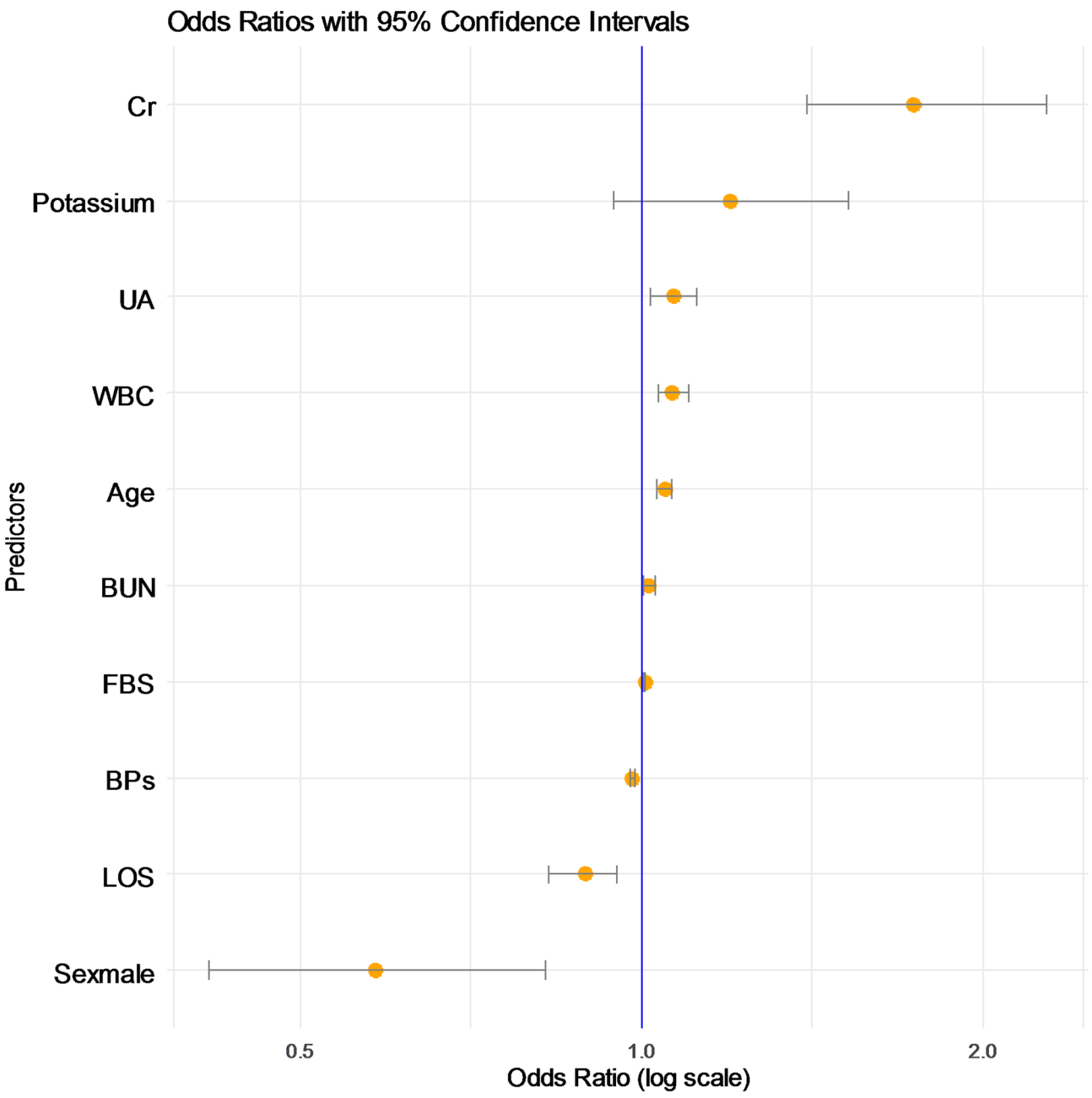

Fig. 8 OR and 95% CI from multivariate regression analysis of selected factors for in-hospital mortality in STEMI

Notwithstanding these limitations, it is hypothesized that the study results may facilitate the screening of STEMI patients at high risk of in-hospital death, thereby contributing to a reduction in preventable mortality. Modifiable risk indicators in the prediction model, other than age and sex, allow medical intervention to improve the assessment of mortality risk in patients.

Conclusion

The risk factor analysis and development of a nomogram for the accurate prediction of death in STEMI patients serves as a valuable tool for clinicians, offering a simplified and precise visualization method for identifying high-risk patients. This, in turn, facilitates a more personalized approach to the management of these cases. The findings from this study contribute not only to the decision-making process for clinicians but also provide a foundation for future research, aiming to enhance

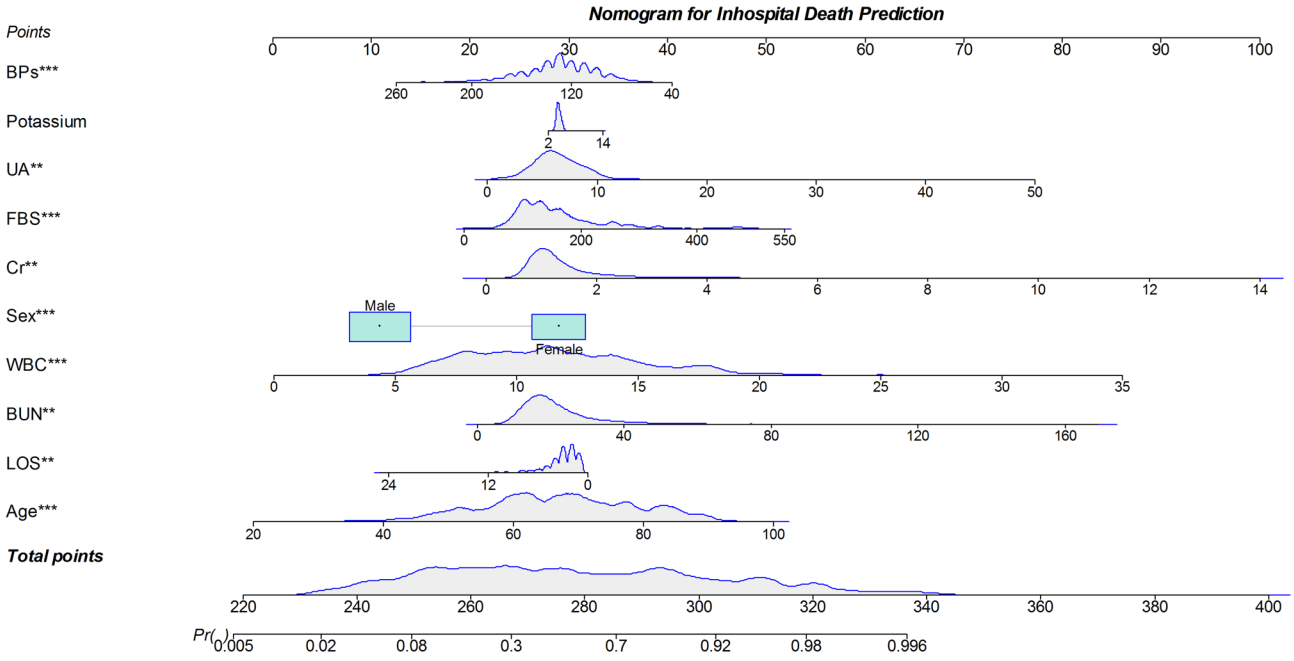

Fig. 9 Nomogram for predicting in-hospital mortality in patients with STEMI and selected risk factors

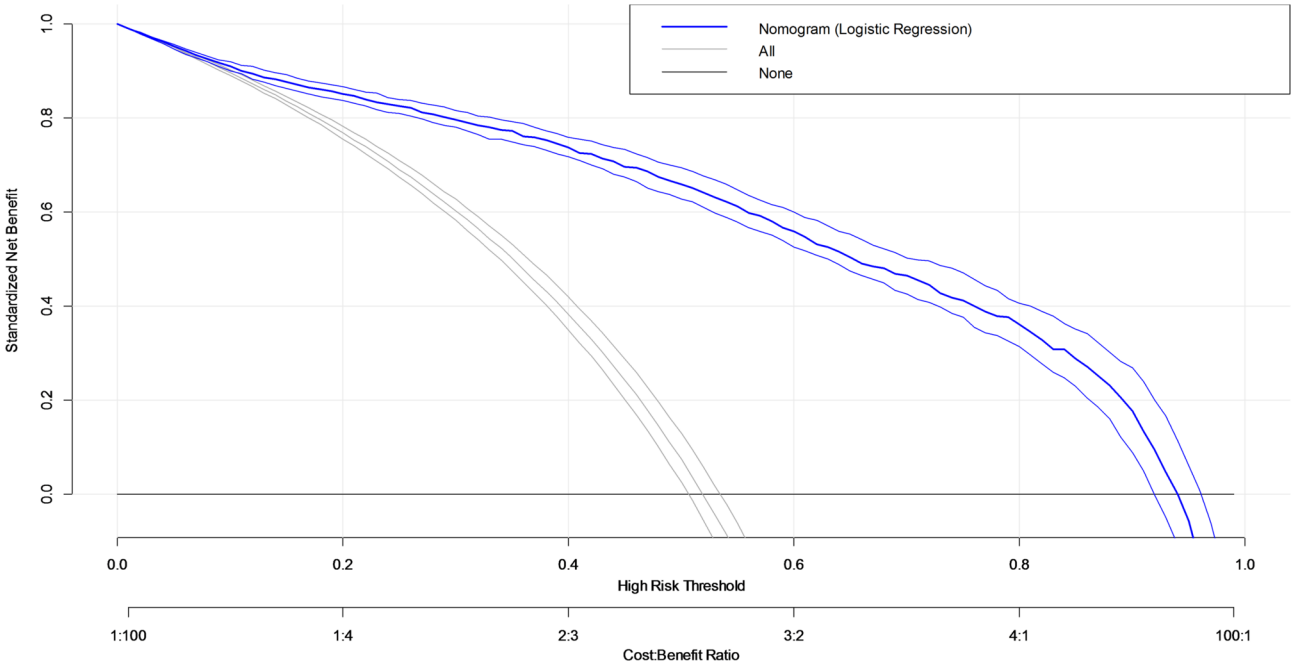

Fig. 10 Decision curve analysis of the nomogram for in-hospital mortality in STEMI patients

the quality of medical care for patients with cardiac conditions.

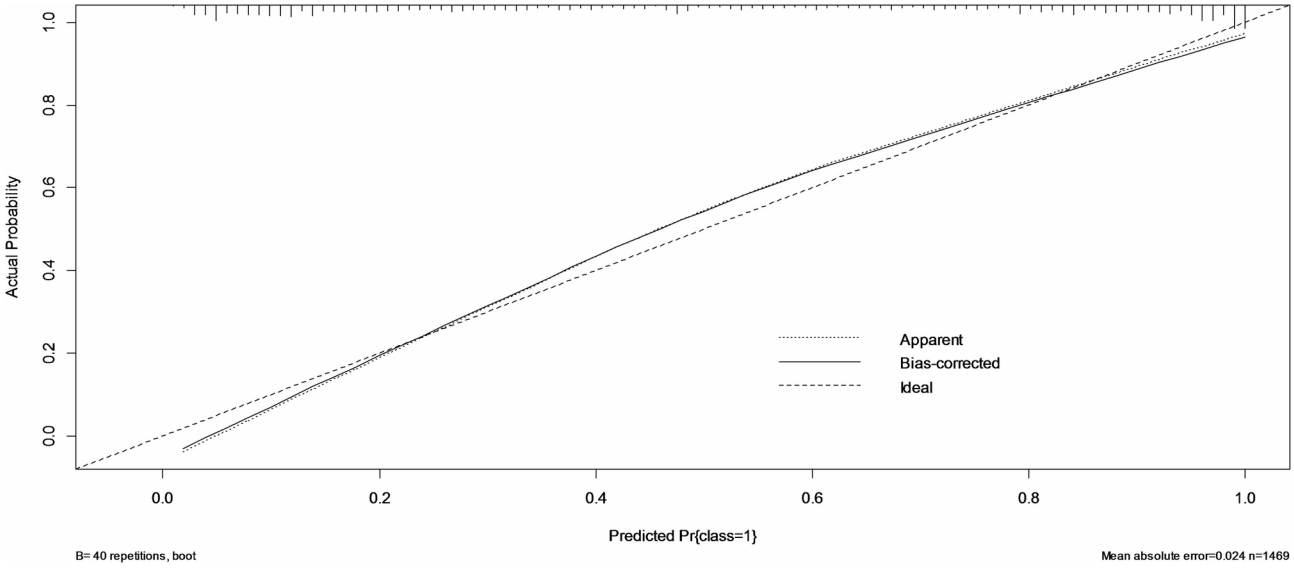

Fig. 11 Calibration plot for in-hospital mortality in STEMI patients

Acknowledgements
This study was supported by Vice-Chancellor of Research and Technology of Hamadan University of Medical Sciences (Grant number: 140208096510). The authors would like to thank the clinical research development unit of Farshchian heart hospital, Hamadan University of medical sciences, Hamadan, Iran for their assistance throughout the period of study.

Author contributions
MF and HM planned the research and conducted the literature review. MF and TR performed data analysis, with MF drafting the initial manuscript. TR and AMY were responsible for data collection. All authors participated in data interpretation, manuscript writing, and approved the final version. MF supervised all aspects of study design, data acquisition, analysis, and article writing.

Funding
This study was supported by Vice-Chancellor of Research and Technology of Hamadan University of Medical Sciences (Grant number: 140208096510).

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

Declarations

Ethics approval and consent to participate
This study was conducted in accordance with the ethical guidelines of the Declaration of Helsinki. Informed consent was waived due to the retrospective design and the data were fully de-identified and no intervention on patients was performed. As the study did not involve any direct interaction with participants and used only anonymised information, it was not necessary to obtain informed consent. The research ethics committee of Hamadan University of Medical Sciences approved the study (IR.UMSHA.REC.1402.349).

Consent for publication
Not applicable.

Relevant guidelines and regulations
This study adhered to all relevant guidelines and regulations for medical research involving human data, including the protection of privacy and confidentiality of personal information.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Biostatistics, School of Public Health, Hamadan University of Medical Sciences, Hamadan, Iran
[2]Research Center for Health Sciences, Institute of Health Sciences and Technologies, Avicenna Health Research Institute, Hamadan University of Medical Sciences, Hamadan, Iran
[3]Department of Cardiology, Faculty of Medicine, Clinical Research Development Unit of Farshchian Hospital, Hamadan University of Medical Sciences, Hamadan, Iran

Received: 8 April 2025 / Accepted: 13 August 2025
Published online: 18 August 2025

References

1. Sarrafzadegan N, Mohammmadifard N. Cardiovascular disease in Iran in the last 40 years: prevalence, mortality, morbidity, challenges and strategies for cardiovascular prevention. Arch Iran Med. 2019;22(4):204–10.
2. McNamara RL, Kennedy KF, Cohen DJ, Diercks DB, Moscucci M, Ramee S, et al. Predicting in-hospital mortality in patients with acute myocardial infarction. J Am Coll Cardiol. 2016;68(6):626–35.
3. Beyranvand MR, Manhoobi H, Shahraz S, Kolahi AA. Myocardial infarction in iran: epidemiology, management, and prognosis. J Tehran Heart Cent. 2023;18(2):82–6.
4. Ahmadi A, Khaledifar A, Sajjadi H, Soori H. Relationship between risk factors and in-hospital mortality due to myocardial infarction by educational level: a National prospective study in Iran. Int J Equity Health. 2014;13:116.
5. Ghadimi H, Bishehsari F, Allameh F, Bozorgi AH, Sodagari N, Karami N, et al. Clinical characteristics, hospital morbidity and mortality, and up to 1-year follow-up events of acute myocardial infarction patients: the first report from Iran. Coron Artery Dis. 2006;17(7):585–91.
6. Gao N, Qi X, Dang Y, Li Y, Wang G, Liu X, et al. Establishment and validation of a risk model for prediction of in-hospital mortality in patients with acute ST-elevation myocardial infarction after primary PCI. BMC Cardiovasc Disord. 2020;20:1–10.
7. Jollis JG, Granger CB, Zègre-Hemsey JK, Henry TD, Goyal A, Tamis-Holland JE, et al. Treatment time and in-hospital mortality among patients with ST-segment elevation myocardial infarction, 2018–2021. JAMA. 2022;328(20):2033–40.
8. Saadatagah S, Ghodsi S, Omidi N, Poorhosseini H, Salarifar M, Sadeghian S, et al. Twelve-Year history of STEMI management in Tehran heart center: concomitant reduction of In-Hospital mortality and hospitalization length. Arch Iran Med. 2020;1(8):514–21.

9. Aziz F, Malek S, Ibrahim KS, Raja Shariff RE, Wan Ahmad WA, Ali RM et al. Short-and long-term mortality prediction after an acute ST-elevation myocardial infarction (STEMI) in asians: a machine learning approach. PLoS ONE 2021;6(8):e0254894.
10. Zhang Z, Geskus RB, Kattan MW, Zhang H, Liu T. Nomogram for survival analysis in the presence of competing risks. Ann Transl Med. 2017;5(20):403.
11. Zhang Z, Kattan MW. Drawing nomograms with R: applications to categorical outcome and survival data. Annals Translational Med. 2017;5(10):211.
12. Hong H, Hong S, simpleNomo:. A python package of making nomograms for visualizable calculation of logistic regression models. Health Data Sci. 2023;3:0023.
13. Jalali A, Alvarez-Iglesias A, Roshan D, Newell J. Visualising statistical models using dynamic nomograms. PLoS ONE. 2019;14(11):e0225253.
14. Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. J Clin Oncol. 2008;26(8):1364–70.
15. Akbar H, Foth C, Kahloon RA et al. Acute ST-Segment Elevation Myocardial Infarction (STEMI) [Updated 2024 Oct 6]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing, Jan 2025 . Available from: https://www.ncbi.nlm.nih.gov/books/NBK532281/
16. Japkowicz N, Stephen S. The class imbalance problem: A systematic study. Intell Data Anal. 2002;6(5):429–49.
17. Ali A, Shamsuddin SM, Ralescu AL. Classification with class imbalance problem. Int J Adv Soft Compu Appl. 2013;5(3):176–204.
18. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. J Stat Softw. 2010;36:1–13.
19. Darst BF, Malecki KC, Engelman CD. Using recursive feature elimination in random forest to account for correlated variables in high dimensional data. BMC Genet. 2018;19:1–6.
20. Muthukrishnan R, Rohini RL. LASSO: A feature selection technique in predictive modeling for machine learning. In. 2016 IEEE international conference on advances in computer applications (ICACA), IEEE 2016:18–20.
21. Speiser JL, Miller ME, Tooze J, Ip E. A comparison of random forest variable selection methods for classification prediction modeling. Expert Syst Appl. 2019;134:93–101.
22. Momeni M, Salari A, Ghanbari A, Rostamnia L. Hospital mortality of patients with acute myocardial infarction and related factor. J North Khorasan Univ Med Sci. 2014;6(2):407–14.
23. Mosa Farkhani E, Baneshi MR, Zolala F. Survival rate and its related factors in patients with acute myocardial infarction. Med J Mashhad Uni Med Sci, 2014;57(4):636–646.
24. Gevaert SA, De Bacquer D, Evrard P, Convens C, Dubois P, Boland J, et al. Gender, TIMI risk score and in-hospital mortality in STEMI patients undergoing primary PCI: results from the Belgian STEMI registry. EuroIntervention. 2014;9(9):1095–101.
25. Zheng X, Dreyer RP, Hu S, Spatz ES, Masoudi FA, Spertus JA, et al. Age-specific gender differences in early mortality following ST-segment elevation myocardial infarction in China. Heart. 2015;101(5):349–55.
26. Bajraktari G, Thaqi K, Pacolli S, Gjoka S, Rexhepaj N, Daullxhiu I, et al. In-hospital mortality following acute myocardial infarction in kosovo: A single center study. Ann Saudi Med. 2008;28(6):430–4.
27. Sanguanwong S, Srimahachota S, Tungsubutra W, Srichaiveth B, Kiatchoosakun S. Predictors of in-hospital mortality in Thai STEMI patients: results from TACSR. J Med Assoc Thai. 2007;90(Suppl 1):91–7.
28. Arslan GY, Çağırıcı G. Indicators of hospital mortality in patients aged 80 years with acute coronary syndrome. Indicators. 2023;11(4):139–46.
29. Swaminathan RV, Rao SV, McCoy LA, Kim LK, Minutello RM, Wong SC, et al. Hospital length of stay and clinical outcomes in older STEMI patients after primary PCI: a report from the National cardiovascular data registry. J Am Coll Cardiol. 2015;65(12):1161–71.
30. Martin L, Murphy M, Scanlon A, Clark D, Farouque O. The impact on long term health outcomes for STEMI patients during a period of process change to reduce door to balloon time. Eur J Cardiovasc Nurs. 2016;15(3):e37–44.
31. Mouhat B, Putot A, Hanon O, Eicher JC, Chagué F, Beer JC et al. Low systolic blood pressure and mortality in elderly patients after acute myocardial infarction. J Am Heart Association 2020;9(5):e013030.
32. Bordejevic DA, Caruntu F, Mornos C, Olariu I, Petrescu L, Tomescu MC et al. Prognostic impact of blood pressure and heart rate at admission on in-hospital mortality after primary percutaneous intervention for acute myocardial infarction with ST-segment elevation in Western Romania. Ther Clin Risk Manag, 2017:1061–8.
33. Adib A, Sepasi N, Fekri Y. Effects of admission-time blood sugar on the intrahospital death of acute myocardial infarction patients. Med Sci J Islamic Azad Univesity-Tehran Med Branch. 2010;20(1):36–9.
34. Barron HV, Harr SD, Radford MJ, Wang Y, Krumholz HM. The the association between white blood cell count and acute myocardial infarction mortality in patients ≥ 65 years of age: findings from the cooperative cardiovascular project. J Am Coll Cardiol. 2001;38(6):1654–61.
35. Jadhav DV, Muley D, Deshmukh S. Effect of white blood cell indices and glycemia on inhospital prognosis of ST-segment elevated myocardial infarction. J Clin Prev Cardiol. 2022;11(1):10–4.
36. Kruk M, Karcz M, Przyłuski J, Bekta P, Kepka C, Kalińczuk L. White blood cell count on admission and mortality in patients treated with primary percutaneous coronary intervention (ANIN myocardial infarction Registry). Postepy Kardiol Interwencyjnej. 2007;3(4):193–4.
37. Liu CF, Song KY, Zhou WN, Wei YJ. Association between uric acid and in-hospital heart failure in patients with acute myocardial infarction undergoing percutaneous coronary intervention. Disease Markers, 2021; 2021(1):7883723.

Publisher's note