Intelligent Medicine 4 (2024) 170–176

Contents lists available at ScienceDirect

Intelligent Medicine

journal homepage: www.elsevier.com/locate/imed

Research Article

Machine learning predicts long-term mortality after acute myocardial infarction using systolic time intervals and routinely collected clinical data

Bijan Roudini [1,2], Boshra Khajehpiri [1], Hamid Abrishami Moghaddam [2], Mohamad Forouzanfar [1,2,3,*]

[1] Department of Systems Engineering, School of Advanced Technology, University of Quebec, Montreal, Canada
[2] Machine Vision and Medical Image Processing (MVMIP) Laboratory, Faculty of Electrical and Computer Engineering, K. N. Toosi University of Technology, Tehran, Iran
[3] Research Center of the Montreal University Institute of Geriatrics, Montreal, Canada

ARTICLE INFO

Keywords:
Electrocardiogram
Predication
Mortality prediction
Machine learning

ABSTRACT

Background Precise estimation of current and future comorbidities of patients with cardiovascular disease is an important factor in prioritizing continuous physiological monitoring and new therapies. Machine learning (ML) models have shown satisfactory performance in short-term mortality prediction in patients with heart disease, whereas their utility in long-term predictions is limited. This study aimed to investigate the performance of tree-based ML models on long-term mortality prediction and effect of two recently introduced biomarkers on long-term mortality.

Methods This study used publicly available data from the Collaboration Center of Health Information Application at the Ministry of Health and Welfare, Taiwan, China. The collected data were from patients admitted to the cardiac care unit for acute myocardial infarction (AMI) between November 2003 and September 2004. We collected and analyzed mortality data up to December 2018. Medical records were used to gather demographic and clinical data, including age, gender, body mass index, percutaneous coronary intervention status, and comorbidities such as hypertension, dyslipidemia, ST-segment elevation myocardial infarction, and non-ST-segment elevation myocardial infarction. Using the data, collected from 139 patients with AMI, from medical and demographic records as well as two recently introduced biomarkers, brachial pre-ejection period (bPEP) and brachial ejection time (bET), we investigated the performance of advanced ensemble tree-based ML algorithms (random forest, AdaBoost, and XGBoost) to predict all-cause mortality within 14 years. A nested cross-validation was performed to evaluate and compare the performance of our developed models precisely with that of the conventional logistic regression (LR) as the baseline method.

Results The developed ML models achieved significantly better performance compared to the baseline LR (C-Statistic, 0.80 for random forest, 0.79 for AdaBoost, and 0.78 for XGBoost, *vs.* 0.77 for LR) ($P_{RF} < 0.001$, $P_{AdaBoost} < 0.001$, and $P_{XGBoost} < 0.05$). Adding bPEP and bET to our feature set significantly improved the performance of the algorithm, leading to an absolute increase in C-statistic of up to 0.03 (C-statistic, 0.83 for random forest, 0.82 for AdaBoost, and 0.80 for XGBoost, *vs.* 0.74 for LR) ($P_{RF} < 0.001$, $P_{AdaBoost} < 0.001$, $P_{XGBoost} < 0.05$).

Conclusion The study indicates that incorporating new biomarkers into advanced ML models may significantly improve long-term mortality prediction in patients with cardiovascular diseases. This advancement may enable better treatment prioritization for high-risk individuals.

1. Introduction

Cardiovascular diseases are the leading cause of death globally, with an estimated 17.9 million deaths per year [1]. Approximately every 40 seconds, a person in the United States experiences myocardial infarction (MI) [2]. In Canada, patients with acute myocardial infarction (AMI) are 4 times more likely to die of all causes in a given year compared to individuals without AMI [3]. In Europe, approximately half of the major coronary events occur in those with a previous hospital discharge of AMI, and 1 out of every 5 patients with AMI suffers a second cardiovascular event in the first year [4]. Given the high risk of cardiovascular events following an MI, patients with MI should be carefully monitored and managed with effective prevention programs.

* Corresponding author: Mohamad Forouzanfar, Department of Systems Engineering, School of Advanced Technology, University of Quebec, Montreal, Canada (Email: mohamad.forouzanfar@etsmtl.ca).

https://doi.org/10.1016/j.imed.2024.01.001
Received 15 January 2023; Received in revised form 10 January 2024; Accepted 16 January 2024

Implantable sensors and smart wearables have provided in-home solutions for continuous monitoring of patients diagnosed with AMI [5,6]. Given the fact that continuous monitoring of patients with AMI is expensive, requires a lot of resources, and is not available to every patient, it is wise to focus on patients with more serious health conditions and those at a higher risk of mortality. A precise estimation of the patient's current and future health situation is important to prioritize them for continuous physiological monitoring and provide them with specific therapies at the earliest possible time [7].

Risk assessment in clinics, particularly for patients with heart disease, often relies on methods such as the global registry of acute coronary events (GRACE) [8] score for patients with acute coronary syndrome (ACS), including patients with AMI, and the thrombolysis in MI (TIMI) [9] score for patients with MI. These scores, based on clinical information and factors such as age and blood pressure, are validated for predicting short-term mortality but may not fully address long-term mortality risks in patients with AMI.

Recently, machine learning (ML) has been employed as a data-driven approach to design accurate classifiers for the risk estimation of patients with cardiovascular diseases [10-19]. ML algorithms can reveal complex hidden patterns in data that cannot be extracted using traditional tools. Artificial neural networks (ANN) and tree-based methods have successfully been used to diagnose heart diseases [20-25] and predict CV outcomes [10-19]. Hernesniemi et al. [13] used XGBoost and logistic regression (LR) to predict the 6-month mortality of patients with ACS. They found that ML models outperformed the GRACE risk score system, with XGBoost achieving the best accuracy. Hadanny et al. [12] showed that the random forest (RF) algorithm performs drastically better than the GRACE score in post-ST-segment elevation MI (post-STEMI) 30-day mortality prediction. Cho et al. [11] reviewed several articles to compare the ability of ML and conventional statistical models to predict readmission and mortality in patients with AMI. Among 19 articles, 13 demonstrated a higher performance for ML models compared to traditional techniques. Lee et al. [16] showed the outperformance of ML models over conventional methods, including GRACE and TIMI scores, in short- and long-term mortality prediction of patients with a history of non-ST-segment elevation MI (NSTEMI). Li et al. [17] developed LR, ANN, and tree-based ML models to predict 1-year post-discharge mortality status of approximately 220,000 patients diagnosed with AMI. The constructed models reached average accuracies of 0.8-0.85, higher than the previously reported in-hospital mortality prediction models. Mansoor et al. [18] studied 10,000 women with STEMI to predict their in-hospital mortality using LR and RF. Their results revealed a comparable performance between the models and confirmed RF as a practical ML tool in clinical settings. Vomlel et al. [19] aimed to predict the 30-day mortality of 603 patients with STEMI using different classifiers, including LR, Bayesian methods, ANN, and tree-based algorithms. Comparing their prediction power, LR and simple Bayesian methods were shown to be the most promising models. Other studies [10,15] with different cohort sizes (5,000 and 22,000 patients) attempted to predict 1-year mortality after the diagnosis of AMI and reported similar risk factors such as age and sex as the most important features.

Most existing studies have focused on short-time mortality prediction and building computational models without recognizing the most important risk factors. Prediction of long-term mortality is essential to prioritize patients who are at higher risk as earlier receivers of medications, preventive interventions, and healthcare plans. In this study, we investigated the application of state-of-the-art tree-based ML models using a set of easy-to-access clinical measures to predict long-term (14 years) mortality in 139 patients (52 survivors and 87 deaths) with AMI. In addition to routinely the collected clinical measures, we considered the recently introduced noninvasively measured systolic time intervals, brachial pre-ejection period (bPEP), and brachial ejection time (bET) [26-28] as novel biomarkers for early prediction of all-cause mortality. The bPEP/bET ratio has proved to be a useful variable for the risk assessment of patients with AMI [29]. In this study, we aimed to investigate its effectiveness on the performance of ML models. We studied the correlation strength between different risk factors and the models' predictive power to find the most important mortality biomarkers in individuals suffering from AMI.

Table 1 Statistics of input data

Variables	Death		*P*-values
	Yes (n = 87)	No (n = 52)	
Age (years, mean±SD)	69.00 ±12.0	55.30 ±11.5	< 0.001
Male (n (%))	60.00 (69.0)	43.00 (83.0)	0.112
Diabetes mellitus (n (%))	21.00 (24.0)	13.00 (25.0)	0.929
Hypertension (n (%))	45.00 (27.0)	15.00 (29.0)	0.014
Dyslipidemia (n (%))	30.00 (34.0)	14.00 (27.0)	0.460
PCI (n (%))	33.00 (38.0)	17.00 (33.0)	0.660
STEMI (n (%))	18.00 (21.0)	12.00 (23.0)	0.906
bPEP (μs, mean±SD)	96.60 ± 21.00	94.80 ± 18.80	0.615
bET (μs, mean±SD)	250.00 ± 34.90	262.40 ± 24.10	0.025
ABI (mean±SD)	0.94 ± 0.20	1.04 ± 0.10	0.001
BMI (kg/m^2, mean±SD)	23.60 ± 3.90	25.70 ± 3.10	0.002

t-test and Chi-square test are performed on numerical and categorical variables, respectively, to examine the dependency between each variable in mortal and nonmortal cases.
PCI: percutaneous coronary intervention; STEMI: ST-segment myocardial infarction; bPEP: brachial pre-ejection period; bET: brachial ejection time; ABI: ankle brachial index; BMI: body mass index.

2. Methods

2.1. Study population

This analysis was based on publicly available data from the Collaboration Center of Health Information Application (CCHIA), Ministry of Health and Welfare, Taiwan, China [29]. The data set comprised patients admitted to the cardiac care unit due to AMI between November 2003 and September 2004. Inclusion criteria were patients older than 20 years diagnosed with type 1 AMI. Exclusion criteria included absence of data on brachial pre-ejection period (bPEP), brachial ejection time (bET), presence of atrial fibrillation, or amputation of extremities. This resulted in a data set of 139 patients (36 women) with AMI aged between 24-91 years. The study tracked and analyzed the mortality of these patients until December 2018, during which 87 patients died. In compliance with *the Declaration of Helsinki*, all patients provided informed consent for their data to be used in the study. Medical records were employed to gather demographic and medical data, including age, gender, body mass index (BMI), percutaneous coronary intervention (PCI) status, and comorbidities such as hypertension, dyslipidemia, STEMI, and NSTEMI. The data from CCHIA were initially collected to assess the potential of the bPEP to bET ratio as a novel biomarker for predicting long-term cardiovascular and overall mortality in patients with AMI [29]. For the purpose of our study, we dichotomized the cohort based on mortality status at the end of the study into 2 groups: those who survived and those who died within the 14-year period post-AMI diagnosis.

2.2. Predictor features

The data set had 11 predictor variables, including clinical and demographic measurements acquired from each participant at baseline, containing 5 numeric and 6 binary features. The numeric variables were bPEP, bET, BMI, ankle-brachial index (ABI), and age. The binary features included PCI received or not received, sex, and medical status of the patients on comorbidities such as dyslipidemia, diabetes, hypertension, and STEMI. ABI, bPEP, and bET were obtained using an ankle-brachial index ABI-form device recorded within 24 h of each participant's admission [29]. Table 1 provides summary statistics of the data measurements at baseline comparing the mortal group to those who survived (nonmortals).

2.3. Machine learning

We selected 3 advanced tree-based ML algorithms to perform our prediction task by finding an optimal decision boundary between the two groups of survivors and mortals. After building our models, they were tested to predict the mortality status of new individuals in the following 14 years.

Tree-based ML methods have proved their strength in modeling complex nonlinear input-output patterns in a nonparametric fashion, especially for structured data [30]. These algorithms achieve the best performance when used in an ensemble learning framework [31]. Ensemble-based ML algorithms work by running multiple base learners and aggregating the decisions made by each to achieve a more reliable outcome. These methods can usually handle noisy data and outliers well with little effect on the overall performance. They are also robust against overfitting and have lower variance than each of their weak learners [32]. Bagging and boosting methods form the two main categories of ensemble learning approaches that are particularly favored over more complex deep learning models in learning tasks where limited data are available.

2.4. Random forest

Random forest (RF) is an ensemble of decision trees that makes predictions by averaging the decisions obtained from all its base learners [33]. RF benefits from 2 sources of randomization: first, the adoption of a random bootstrap of the data for each tree building, and second, restricting the candidate features to a random subset at each node splitting level of the trees. Averaging the results of multiple models along with the randomization helps the forest to maintain a low generalization error while growing deeper trees.

2.5. Boosting machines

In the realm of advanced boosting techniques, adaptive boosting (AdaBoost) [34] and extreme gradient boosting (XGBoost) [35] stand out as notable extensions of the traditional boosting methods. AdaBoost is characterized by its sequential ensemble model structure and minimal hyperparameter configuration, which lends it robustness against overfitting, particularly in scenarios involving low-noise and smaller data sets. However, XGBoost represents an efficient iteration of the gradient tree boosting methodology. Its primary strengths lie in its rapid processing speed, surpassing other similar algorithms, and its regularization feature, which significantly reduces variance and enhances model performance.

2.6. Data analysis

The data set we used did not contain any missing values. One-hot encoding was applied to convert the categorical features into binary variables [36]. To enhance the numerical stability of the models, standard scaling (z-score normalization) was applied by subtracting the mean from the feature values and then dividing them by the standard deviation [37].

Model selection, hyperparameter optimization, and performance evaluation were performed using nested cross-validation to verify the generalization ability of the ML models. This data division technique helps to estimate an unbiased generalization performance of a model and is specifically useful when dealing with smaller data sizes. The cross-validation method was implemented as follows: A 10-fold outer loop was first formed by dividing the whole data randomly into ten equal parts. For each fold, 90% of the samples entered a 5-fold inner loop for model selection and hyperparameter tuning, while the remaining 10% was set aside for final testing. This procedure was repeated 10 times to reduce the inherent randomness of ML models, and the average results were reported. Notably, all the reported results are on the test set, which was not seen during training and validation.

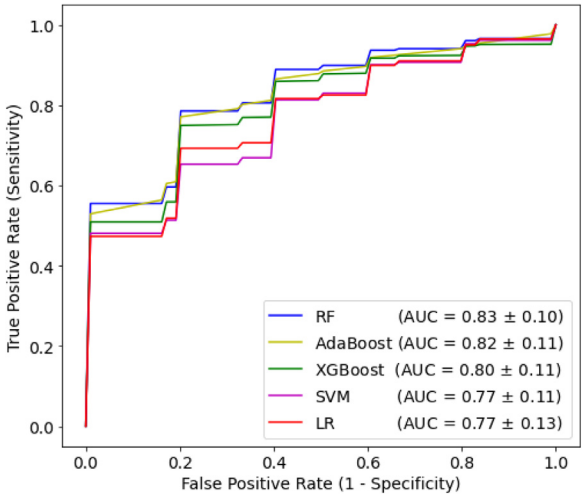

Figure 1. Receiving operating characteristics curves for different acute myocardial infarction mortality prediction models. RF: random forest, AdaBoost: adaptive boosting, XGBoost: extreme gradient boosting, SVM: support vector machine, LR: logistic regression.

Table 2 lists the hyperparameters, their search ranges, and their optimal values for each ML model. They include the general learning hyperparameters of the models (e.g., learning rate, input sampling method, and node splitting criteria) as well as those that determine model complexity and control underfitting/overfitting (e.g., number of estimators, maximum depth of the trees, pruning thresholds, and feature sampling rate at each node).

2.7. Statistical analysis

Our designed ML models were evaluated based on the receiving operating characteristics (ROC) metric. The models were compared with each other using the area under ROC curve (AUC) index, also known as C-statistic, accuracy, sensitivity (true positive rate, recall), specificity, and precision.

Paired *t*-tests were performed on the cross-validation results to investigate whether the performances of different ML methods are superior to those of the baseline LR. A one-way repeated measure, analysis of variance (ANOVA) was performed to compare the performances of the three ML models, followed by Tukey post hoc multiple comparisons.

The Gini feature importance method [38] was used to rank the most important predictive features. To investigate the effect of bPEP and bET on the prediction performance, the same experiments were repeated by excluding the two parameters from the analysis.

3. Results

3.1. Cohort characteristics

The study longitudinally followed 139 patients initially diagnosed with AMI. Among these, 87 participants succumbed to death, and 52 survived until the final follow-up.

3.2. Model performance

As depicted in Figure 1, the average ROC curves demonstrated the performance of the developed models. Table 3 details their predictive capabilities. Notably, the RF model exhibited the highest AUC of 0.83. AdaBoost excelled in accuracy (82%), sensitivity (90%), specificity (69%), and precision (79%), surpassing other models in these metrics.

Table 2 Hyperparameter tuning for ML algorithms

Models	Hyperparameters	Search interval	Optimum values
RF	Number of estimators	(20, 50, 150, 200, 250, 300, 350, 400, 500, 1000)	250
	Max number of features	("auto", "sqrt")	"sqrt"
	Max depth	(1, 3, 5, 6, 7, 12, 14, 16, 18)	3
	Min samples split	(2, 4, 6, 8, 10, 12, 14)	10
	Min samples leaf	(2, 3, 4, 5, 6, 7, 9, 12)	6
	Bootstrap	(True, False)	False
	Criterion	("entropy", "gini")	"entropy"
AdaBoost	Number of estimators	(20, 50, 100, 300, 400, 500, 1000)	20
	Learning rate	(0.001, 0.01, 0.05, 0.1, 0.5)	0.01
XGBoost	Number of estimators	(20, 50, 100, 300, 400, 500, 1000)	100
	Max depth	(1, 3, 5, 6, 7, 10)	1
	Eta	(0.01, 0.03, 0.05, 0.1, 0.2)	0.1
	Min child weight	(0.1, 0.3, 0.5)	0.3
	Max leaf nodes	(4, 6, 9, 10)	9
	Subsample	(0.1, 0.5, 0.8, 1)	0.5
	Gamma	(0.01, 0.05, 0.1, 0.2, 0.5, 0.8)	0.1
	Alpha	(0.001, 0.01, 0.1, 0.5)	0.5
	Max delta step	(0, 1, 2, 5)	2
	Colsample by tree	(0.5, 0.6, 0.8)	0.5
	Colsample by level	(0.4, 0.6, 0.8)	0.6
	Colsample by node	(0.2, 0.3, 0.5, 0.8)	0.3
	Lambda	(0.01, 0.05, 0.1, 0.3, 0.5)	0.05

RF: random forest, AdaBoost: adaptive boosting, XGBoost: extreme gradient boosting, ML: machine learning

Table 3 MI mortality prediction performance of ensemble tree-based ML models

Algorithms	C-Statistic (AUC)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
RF	0.83 ± 0.10***	77 ± 9***	88 ± 11***	60 ± 20*	79 ± 9**
AdaBoost	0.82 ± 0.11***	78 ± 10***	85 ± 13***	67 ± 20***	82 ± 9***
XGBoost	0.80 ± 0.11*	76 ± 9***	82 ± 13*	67 ± 21***	81 ± 10***
LR	0.77 ± 0.13	71 ± 10	81 ± 12	56 ± 23	76 ± 10

P values were obtained using paired *t*-test in comparison to the baseline LR mode.
RF: random forest, AdaBoost: adaptive boosting, XGBoost: extreme gradient boosting, LR: logistic regression, ML: machine learning.
* Implies *P* values < 0.05.
** Implies *P* values < 0.01.
*** Implies *P* values < 0.001.

Table 4 Effect of bPEP and bET on the prediction performance of different ML models

Experiments	Algorithm	C-statistic (AUC)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
Experiment I (with all features)	RF	0.83***	77***	88**	60***	79***
	AdaBoost	0.82***	78***	85***	67	82
	XGBoost	0.80*	76	82	67	81
Experiment II (bET & bPEP excluded)	RF	0.80	73	86	52	75
	AdaBoost	0.79	78	85	67	81
	XGBoost	0.78	77	83	67	81

RF: random forest, AdaBoost: adaptive boosting, XGBoost: extreme gradient boosting, ML: machine learning.
* Implies *P* values < 0.05.
** Implies *P* values < 0.01.
*** Implies *P* values < 0.001.

Compared to the baseline LR, all ML models showed significant enhancements across all classification metrics.

3.3. Variable importance

Figure 2 presents the relative importance of the input variables in RF, AdaBoost, and XGBoost models. Age, BM, ABI, bET, and bPEP were identified as the most influential factors across all models. Conversely, variables such as diabetes, dyslipidemia, PCI, STEMI, and sex demonstrated the least predictive power.

3.4. Effect of bPEP and bET

Our analysis focused on assessing the impact of including bPEP and bET in the predictive models. This involved retraining and testing the algorithms with and without these parameters, as detailed in Table 4 and Supplementary Tables 1 and 2.

In Experiment 1, which included all the features, the RF model exhibited superior performance with an AUC of 0.83, an accuracy of 77%, sensitivity of 88%, specificity of 60%, and precision of 79% (all with $P < 0.001$). AdaBoost also performed well with an AUC of 0.82, accuracy of 78%, and precision of 82%, and XGBoost had an AUC of 0.80. The training results for Experiment 1, as shown in Supplementary Table 1, mirrored these findings, with RF and AdaBoost demonstrating high AUCs of (0.83 ± 0.11) and (0.82 ± 0.10), respectively, and XGBoost showing a slightly lower AUC of (0.81 ± 0.12).

In contrast, Experiment 2, where bPEP and bET were excluded, showed a noticeable decline in the performance metrics for all models. The RF model's AUC decreased to 0.80, with the corresponding drops in accuracy, sensitivity, specificity, and precision. AdaBoost and XGBoost also showed reduced performance, as reflected in their AUCs of 0.79

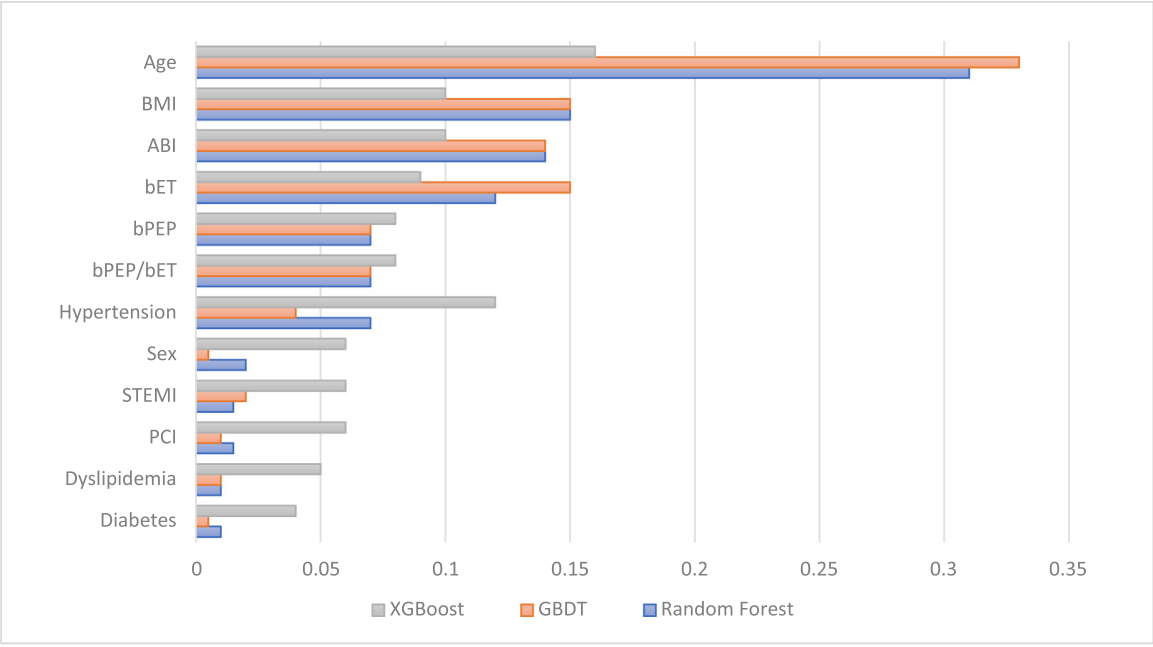

Figure 2. Variable importance for AMI mortality prediction. AMI: acute myocardial infarction.

and 0.78, respectively. The training results for Experiment 2, detailed in Supplementary Table 2, consistently revealed lower performance metrics across all models compared to Experiment 1.

The inclusion of bPEP and bET significantly improved the RF model's performance across all classification metrics (P < (0.001-0.01)). AdaBoost showed marked improvements in AUC, accuracy, and sensitivity (P < 0.001), and XGBoost demonstrated a significant increase in AUC (P < 0.05) when these parameters were included. These results underscore the importance of bPEP and bET as predictive features in the models, contributing substantially to their accuracy and reliability in predicting mortality in patients with AMI.

4. Discussion

In this study, tree-based ML algorithms were developed for the mortality prediction of patients diagnosed with AMI over a period of 14 years. Our models received clinical and demographic information as well as two noninvasively measured systolic time intervals, bPEP and bET. Among the different ML algorithms, ensemble tree-based ML algorithms were selected as they perform drastically better than other algorithms in most applications with limited available data [31]. Bagging (RF) and two boosting (AdaBoost and XGBoost) ensemble algorithms were considered in this study.

It was observed that all the ML models significantly outperformed the baseline LR algorithm. The best results were achieved using RF and AdaBoost, where RF achieved the highest C-statistic (0.83) and sensitivity (88%), and AdaBoost achieved the highest accuracy (47%), specificity (67%), and precision (82%). The AdaBoost classifier correctly predicted 78 mortal patients out of 87 mortal cases, indicating that with a spanning of 70% of the study population, approximately 90% of mortal cases were distinguished. RF closely followed AdaBoost by correctly predicting 77 mortal patients out of the 87 mortal cases. XGBoost slightly underperformed compared to RF and AdaBoost.

Additionally, it is crucial to highlight the consistency in model performance across training and testing sets, as indicated in Supplementary Tables 1 and 2 for Experiments 1 and 2. The minor differences observed between these sets underscore the reliability and generalizability of the models, suggesting their robustness in practical applications.

The one-way ANOVA test was performed to find differences in the performances of the ML classifier. A meaningful difference in sensitivity and specificity between ML models was observed (P < 0.05). The Tukey multiple comparison test was performed to find the superior models. It was observed that RF performed significantly better than XGBoost in terms of sensitivity, whereas AdaBoost and XGBoost performed significantly better than RF in terms of specificity. Although AdaBoost achieved higher mean sensitivity compared to RF (88% *vs*. 85%), it attained a higher standard deviation of sensitivity (13% *vs*. 10%). A higher sensitivity indicates a lower false negative rate, which is extremely important in the prediction of mortality in patients with AMI. Given that there was no significant difference between RF and AdaBoost models in terms of sensitivity, both algorithms can be considered equally suitable for our application. In terms of model complexity, AdaBoost is superior to RF as it only requires 20 estimators compared to 250 estimators for RF.

One of the objectives of this study was to examine the usefulness of the recently introduced features, bPEP and bET, by using ML models in the prediction of long-term mortality of patients with AMI. Pre-ejection preiod is the time taken from the electrical depolarization of the left ventricle to the beginning of ventricular ejection, i.e., when the aortic valve opens. It is an index of myocardial contractility and beta-adrenergic sympathetic control of the heart. Ejection time is the time taken from the opening of the aortic valve to its closure and is conventionally used to evaluate the ventricle function and contractibility. Heart impairments usually prolong pre-ejection period and shorten ejection time [39,40].

The higher values of bPEP/bET are shown to have a high correlation with cardiovascular mortality [29]. bPEP and bET can be easily calculated from the morphology of noninvasively measured pulse waveform and electrogardiogram; therefore, they can be easily integrated into any predictive model with minimum cost. According to Table 4, adding bPEP and bET to our input feature set led to the highest prediction performance, which implies their importance as new biomarkers of mortality in patients with AMI. Moreover, by considering bPEP and bET, the RF model's AUC, accuracy, sensitivity, specificity, and precision improved by 3%, 4%, 2%, 8%, and 4%, respectively. For the AdaBoost model, AUC and precision improved by 3% and 1%, respectively, whereas other classification metrics remained unchanged. It can, therefore, be concluded that RF can better extract the additional predictive information from

the new features owing to its more advanced architecture with a greater number of estimators.

Analyzing the importance of features in an ML model is another way of finding the most useful features for mortality prediction. Age, ABI, BMI, bET, bPEP, and hypertension were among the most important factors for mortality prediction. Most of these factors also had significant differences between mortal and nonmortal cases (Table 1). Our findings are in accordance with the results of previous studies that indicated that age was an important factor in predicting the overall mortality of patients after AMI and heart failure [15,41].

This study was limited to the application of tree-based algorithms for the prediction of long-term mortality of patients with AMI. ANN models, especially those with deep architectures, have recently dominated the computational biology field. ANN-based algorithms can perform better than other ML models when a large volume of training data are available. This study was limited to the analysis of data collected from 139 individuals, and therefore, tree-based algorithms were selected to avoid overfitting and achieve a more generalizable assessment. Future work should be directed toward the collection of a larger and more detailed data set where advanced deep learning models can be applied to predict a more precise time of mortality in the future. Another limitation of this study was the unbalanced sex distribution (36 women out of 139) while being a woman diagnosed with AMI is an important risk factor [42]. By collecting a balanced data set on all genders, the effect of sex on the mortality of patients diagnosed with AMI can further be studied.

In clinical practice, evidence-based risk assessment methods are integral to distinguishing heart disease patients requiring intensive care unit level attention. Prominent among these is the GRACE score, a widely used tool for stratifying risk in patients with ACS [8]. Given that AMI is a subset of ACS, the GRACE score becomes pertinent for assessing the mortality/MI risk post-AMI. This score is derived from various clinical parameters, including age, heart rate, and systolic blood pressure. Its efficacy in predicting short-term (6 months to 1 year) all-cause mortality for AMI patients is well-established [43-47]. Another method, the TIMI score, is employed to estimate short-term mortality in patients with MI, incorporating factors such as age, coronary artery disease (CAD) risk factors, and clinical history [9]. GRACE and TIMI scores are thus pivotal in assessing short-term mortality risks. Although our study's tree-based algorithms offer insightful data, they are not directly comparable to traditional GRACE and TIMI scores due to the lack of certain factors in our data set necessary for these models. Nevertheless, future studies could explore a comparative analysis between these advanced tree-based algorithms and traditional models such as GRACE and TIMI.

Notably, the nested cross-validation method used in this study, although beneficial in reducing model overfitting, introduces minor dependencies between training and test sets. This interdependency slightly contravenes the paired t-test assumption of independence between folds, presenting a methodological limitation [48,49]. Despite this, the impact on our results is expected to be minimal, owing to the rigorous nature of nested cross-validation [50].

5. Conclusion

This study investigated the all-cause long-term mortality prediction power of different ensemble tree-based ML algorithms using routinely collected clinical data as well as two new biomarkers, bPEP and bET. It was found that ML models can accurately predict mortality with a C-statistic as high as 0.83, which is statistically superior to LR as the baseline model. We also demonstrated that adding bPEP and bET to our input feature set improves the prediction results in most of the evaluation metrics. Predicting the mortality status of patients diagnosed with AMI over a long period of 14 years after AMI was another critical feature of this study. Our fast and easy-to-use ML models can assist medical staff in prioritizing patients with AMI for intense monitoring and preventing severe outcomes. Future research should be directed toward enhancing the prediction performance by collecting a larger balanced data set and using more advanced ML algorithms such as ANNs.

Conflicts of interest statement

The authors have no conflicts of interest to declare.

Funding

The authors received no financial support for this research.

Author contributions

Bijan Roudini, implemented the algorithms, performed the analysis, and wrote the initial draft. Boshra Khajehpiri checked and evaluated the results, contributed to generating ideas, and assisted in writing the draft. Professor Hamid Abrishami Moghaddam and Professor Mohamad Forouzanfar jointly supervised the project, participated in generating the core ideas, and extensively edited and revised the draft.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.imed.2024.01.001.

References

[1] World Health Organization. Cardiovascular diseases (CVDs). 2021. Available from: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (Accessed on 9 October 2022).
[2] Table 13-10-0394-01 Leading causes of death, total population, by age group. 2021. Available from: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310039401.
[3] Report from the Canadian Chronic Disease Surveillance System: Heart Disease in Canada, 2018. Available from: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/report-heart-disease-Canada-2018.html.
[4] Jernberg T, Hasvold P, Henriksson M, et al. Cardiovascular risk in post-myocardial infarction patients: Nationwide real world data demonstrate the importance of a long-term perspective. Eur Heart J 2015;36(19):1163–70. doi:10.1093/eurheartj/ehu505.
[5] Bayoumy K, Gaber M, Elshafeey A, et al. Smart wearable devices in cardiovascular care: Where we are and how to move forward. Nat Rev Cardiol 2021;18(8):581–99. doi:10.1038/s41569-021-00522-7.
[6] Jons C, Thomsen PE. Treating cardiac arrhythmias detected with an implantable cardiac monitor in patients after an acute myocardial infarction. Curr Treat Options Cardiovasc Med 2012;14(1):39–49. doi:10.1007/s11936-011-0163-4.
[7] Albahri OS, Albahri AS, Mohammed KI, et al. Systematic review of real-time remote health monitoring system in triage and priority-based sensor technology: Taxonomy, open challenges, motivation and recommendations. J Med Syst 2018;42(5):80. doi:10.1007/s10916-018-0943-4.
[8] Tang EW, Wong CK, Herbison P. Global Registry of Acute Coronary Events (GRACE) hospital discharge risk score accurately predicts long-term mortality post acute coronary syndrome. Am Heart J 2007;153(1):29–35. doi:10.1016/j.ahj.2006.10.004.
[9] Bradshaw PJ, Ko DT, Newman AM, et al. Validation of the Thrombolysis In Myocardial Infarction (TIMI) risk index for predicting early mortality in a population-based cohort of STEMI and non-STEMI patients. Can J Cardiol 2007;23(1):51–6. doi:10.1016/s0828-282x(07)70213-1.
[10] Barrett LA, Payrovnaziri SN, Bian J, et al. Building computational models to predict one-year mortality in ICU patients with acute myocardial infarction and post myocardial infarction syndrome. AMIA Jt Summits Transl Sci Proc 2019;2019:407–16.
[11] Cho SM, Austin PC, Ross HJ, et al. Machine learning compared with conventional statistical models for predicting myocardial infarction readmission and mortality: A systematic review. Can J Cardiol 2021;37(8):1207–14. doi:10.1016/j.cjca.2021.02.020.
[12] Hadanny A, Shouval R, Wu J, et al. Predicting 30-day mortality after ST elevation myocardial infarction: Machine learning- based random forest and its external validation using two independent nationwide datasets. J Cardiol 2021;78(5):439–46. doi:10.1016/j.jjcc.2021.06.002.
[13] Hernesniemi JA, Mahdiani S, Tynkkynen JA, et al. Extensive phenotype data and machine learning in prediction of mortality in acute coronary syndrome - The MADDEC study. Ann Med 2019;51(2):156–63. doi:10.1080/07853890.2019.1596302.
[14] Khera R, Haimovich J, Hurley NC, et al. Use of machine learning models to predict death after acute myocardial infarction. JAMA Cardiol 2021;6(6):633–41. doi:10.1001/jamacardio.2021.0122.
[15] Lee HC, Park JS, Choe JC, et al. Prediction of 1-year mortality from acute myocardial infarction using machine learning. Am J Cardiol 2020;133:23–31. doi:10.1016/j.amjcard.2020.07.048.

[16] Lee W, Lee J, Woo SI, et al. Machine learning enhances the performance of short and long-term mortality prediction model in non-ST-segment elevation myocardial infarction. Sci Rep 2021;11(1):12886. doi:10.1038/s41598-021-92362-1.
[17] Li YM, Jiang LC, He JJ, et al. Machine learning to predict the 1-year mortality rate after acute anterior myocardial infarction in Chinese patients. Ther Clin Risk Manag 2020;16:1–6. doi:10.2147/TCRM.S236498.
[18] Mansoor H, Elgendy IY, Segal R, et al. Risk prediction model for in-hospital mortality in women with ST-elevation myocardial infarction: A machine learning approach. Heart Lung 2017;46(6):405–11. doi:10.1016/j.hrtlng.2017.09.003.
[19] Vomlel J, Kruzık H, Tuma P, et al. Machine learning methods for mortality prediction in patients with ST elevation myocardial infarction. In: Proceedings of WUPES 2012; 2012. p. 204–13.
[20] Elhoseny M, Abed Mohammed MA, Mostafa S, et al. A new multi-agent feature wrapper machine learning approach for heart disease diagnosis. Comput Mater Contin 2021;67(1):51–71. doi:10.32604/cmc.2021.012632.
[21] Maji S, Arora S. Decision tree algorithms for prediction of heart disease. Information and communication technology for competitive strategies 2018:447–54. doi:10.1007/978-981-13-0586-3_45.
[22] Malav A, Kadam K. A hybrid approach for heart disease prediction using artificial neural network and K-means. Int J Pure Appl Math 2018;118(8):103–10.
[23] Pathak AK, Arul Valan J. A predictive model for heart disease diagnosis using fuzzy logic and decision tree. Adv Intell Syst Comput 2020:131–40. doi:10.1007/978-981-13-9680-9_10.
[24] Saxena K, Sharma R. Efficient heart disease prediction system using decision tree. In: Proceedings of International Conference on Computing, Communication & Automation. IEEE; 2015.
[25] Shouman M, Turner T, Stocker R. Using decision tree for diagnosing heart disease patients. Aust Comput Soc 2011.
[26] Chen SC, Chang JM, Tsai JC, et al. A new systolic parameter defined as the ratio of brachial pre-ejection period to brachial ejection time predicts overall and cardiovascular mortality in hemodialysis patients. Hypertens Res 2010;33(5):492–8. doi:10.1038/hr.2010.24.
[27] Chen SC, Chang JM, Tsai JC, et al. A systolic parameter defined as the ratio of brachial pre-ejection period to brachial ejection time predicts cardiovascular events in patients with chronic kidney disease. Circ J 2010;74(10):2206–10. doi:10.1253/circj.cj-10-0273.
[28] Su HM, Lin TH, Lee CS, et al. Myocardial performance index derived from brachial-ankle pulse wave velocity: A novel and feasible parameter in evaluation of cardiac performance. Am J Hypertens 2009;22(8):871–6. doi:10.1038/ajh.2009.94.
[29] Su HM, Lee WH, Tsai WC, et al. Usefulness of the ratio of brachial pre-ejection period to brachial ejection time in prediction of cardiovascular and overall mortality in patients with acute myocardial infarction. PLOS ONE 2021;16(1):e0245860. doi:10.1371/journal.pone.0245860.
[30] Wang J, Hu B, Li X, et al. GTC Forest: An ensemble method for network structured data classification. In: Proceedings of 2018 14th International Conference on Mobile Ad-Hoc and Sensor Networks (MSN); 2018. doi:10.1109/msn.2018.00020.
[31] Zhou ZH. Ensemble learning. Springer; 2021.
[32] Dietterich TG. Ensemble learning. The handbook of brain theory and neural networks 2002;2(1):110–25.
[33] Breiman L. Machine learning 2001;45(1):5–32. doi:10.1023/a:1010933404324.
[34] Schapire RE. Explaining AdaBoost. Empirical Inference 2013:37–52. doi:10.1007/978-3-642-41136-6_5.
[35] Chen T, Guestrin C. XGBoost. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016. doi:10.1145/2939672.2939785.
[36] Budholiya K, Shrivastava SK, Sharma V. An optimized XGBoost based diagnostic system for effective prediction of heart disease. J King Saud Univ Comput Inf Sci 2022;34(7):4514–23. doi:10.1016/j.jksuci.2020.10.013.
[37] Patro SG, Sahu KK. Normalization: A preprocessing stage. Int Adv Res J Sci Eng Technol 2015:20–2. doi:10.17148/iarjset.2015.2305.
[38] Nembrini S, König IR, Wright MN. The revival of the Gini importance? Bioinformatics 2018;34(21):3711–18. doi:10.1093/bioinformatics/bty373.
[39] Garrard CL Jr, Weissler AM, Dodge HT. The relationship of alterations in systolic time intervals to ejection fraction in patients with cardiac disease. Circulation 1970;42(3):455–62. doi:10.1161/01.cir.42.3.455.
[40] Tavakolian K. Systolic time intervals and new measurement methods. Cardiovasc Eng Technol 2016;7(2):118–25. doi:10.1007/s13239-016-0262-1.
[41] Chicco D, Jurman G. Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. BMC Med Inform Decis Mak 2020;20(1):16. doi:10.1186/s12911-020-1023-5.
[42] McNair PW, Bilchick KC, Keeley EC. Very late presentation in ST elevation myocardial infarction: Predictors and long-term mortality. Int J Cardiol Heart Vasc 2019;22:156–9. doi:10.1016/j.ijcha.2019.02.002.
[43] Fox KA, Dabbous OH, Goldberg RJ, et al. Prediction of risk of death and myocardial infarction in the six months after presentation with acute coronary syndrome: Prospective multinational observational study (GRACE). BMJ 2006;333(7578):1091. doi:10.1136/bmj.38985.646481.55.
[44] Hung J, Roos A, Kadesjö E, et al. Performance of the GRACE 2.0 score in patients with type 1 and type 2 myocardial infarction. Eur Heart J 2021;42(26):2552–61. doi:10.1093/eurheartj/ehaa375.
[45] Koonsiripaiboon E, Tungsubutra W. Validation of the GRACE risk score to predict in-hospital mortality in patients with ST segment elevation myocardial infarction in Thailand. J Med Assoc Thai 2013;96(Suppl 2):S139–45.
[46] Luo JG, Yang M, Han L, et al. Validity of the global registry of acute coronary events risk score in prediction of acute myocardial infarction mortality in hospitalised Chinese patients aged 80 and over. Australas J Ageing 2014;33(4):E1–5. doi:10.1111/ajag.12044.
[47] Neves VB, Roman RM, Vendruscolo T, et al. Validation of the Grace risk score to predict in-hospital and 6-month post-discharge mortality in patients with acute coronary syndrome. Int J Cardiovasc Sci 2021. doi:10.36660/ijcs.20210019.
[48] Grandvalet Y, Bengio Y. Hypothesis testing for cross-validation. Montreal Universite de Montreal; 2006.
[49] Wong TT, Yang NY. Dependency analysis of accuracy estimates in k-fold cross validation. IEEE Trans Knowl Data Eng 2017;29(11):2417–27. doi:10.1109/tkde.2017.2740926.
[50] Wong TT, Yeh PY. Reliable accuracy estimates from k-fold cross validation. IEEE Trans Knowl Data Eng 2020;32(8):1586–94. doi:10.1109/tkde.2019.2912815.