Machine Learning to Predict 30 Days and 1-Year Mortality in STEMI and Turndown Patients

Aleeha Iftikhar[2], Raymond R Bond[2], Victoria McGilligan[5], Stephen J Leslie[3], Charles Knoery[3], Khaled Rjoob[2], Anne McShane[4], Aaron Peace[1]

[1]Department of Cardiology, Altnagelvin Hospital, Western Health and Social Care Trust, UK
[2]Ulster University, Computing Engineering and Build Environment, Northern Ireland, UK
[3]Cardiac Unit, Raigmore Hospital, Inverness, UK
[4]Letterkenny University Hospital, Letterkenny, Ireland
[5]Ulster University, Centre for Personalised Medicine, Northern Ireland, UK

Abstract

Primary percutaneous coronary intervention (PPCI) is a minimally invasive procedure to unblock the arteries which carry blood to the heart. Referred patients are accepted or turned down for PPCI mainly based on the presence of ST segment elevation on the surface electrocardiogram. We explored the features which predict 30 days and 1-year mortality in accepted and turndown patients and report the performance of machine learning (ML) algorithms. Different ML algorithms, namely multiple logistic regression (MLR), decision tree (DT), and a support vector machine (SVM) were used for the prediction of 30 days and 1-year mortality. Upon significance of various features to predict the 30 days and 1-year mortality, the accuracy, sensitivity, and specificity were compared between algorithms. DT outperformed the other algorithms (SVM and MLR) to predict mortality of patients referred to the PPCI service. Greater sensitivity is achieved in predicting 30 days mortality in the accepted group compared to the turndown group, however, the former model included more features.

1. Introduction

Primary percutaneous coronary intervention (PPCI) is a surgical treatment to unblock arteries that carry blood to the heart. Before conducting PPCI, the 12-lead electrocardiogram (ECG) is often used with ST-elevation myocardial infarction (STEMI) criteria to diagnose and refer patients to a PPCI service. When the patient enters the hospital to the time when a balloon-tipped catheter via the PPCI service is used to reopen the blocked artery is referred to as the 'door-to-balloon time' (DTBT). DTBT is a key performance indicator for the PPCI service. Scholz et al. [1] investigated and concluded that there is no consequence of contact-to-balloon time on mortality in STEMI patients with and without hemodynamic instability. Also mortality in relation to sex and race alone resulting in no statistically significant findings [2]. ML has been previously used to predict mortality. The feasibility and accuracy of ML to predict 5-year all-cause mortality (ACM) in patients undergoing coronary computed tomographic angiography (CCTA) concluded that ML combining clinical and CCTA data was found to predict 5-year ACM significantly better than existing clinical or CCTA metrics alone [3]. The research has also shown that ML can be used for outcome prediction in complex cardiology settings [4]. One study found that a support vector machine (SVM) based on different combinations of heart-rate dynamics and heart-rate variability (HRV) yielded the largest AUC and was the best classifier for predicting cardiac death after AMI [5]. Different ML models using various electronic health records data variables to predict all-cause mortality. It was observed that ML approaches applied to raw EHR data can be used to build models for use in research and clinical practice, and identify novel predictive variables and their effects to inform future research [6]. The literature shows that there have been studies on predicitng mortality, however, there is still a lack of evidence to unveil the factors that could possibly affect or predict 30 days and 1-year mortality in patients referred to PPCI. The aim of this research is to analyse a PPCI dataset containing referred patients who were accepted and turned down for PPCI to discover the features/factors contributing to predicting 30 days and 1-year mortality.

2. Methods

A retrospective analysis of the PPCI dataset was conducted. This dataset recorded the patient's data with a suspected STEMI that were accepted or turned down

between September 2014 to February 2019 at Altnagelvin hospital (Northern Ireland, UK). For missing values, data imputation was performed using Multivariate Imputation by Chained Equations (MICE package in R). Keeping the dominant factors of the dataset in mind, these factors were included in the prediction analysis. All statistical analyses were performed using R-Studio and the R programming language. The visualisations were generated using ggplot2 (an R package for visual analytics).

Multiple feature selection algorithms are applied to elicit variables contributing to 30 days and 1-year mortality for the accepted and turndown patients. Feature selection techniques included correlation, Logistic regression, Minimum redundancy and maximum relevance (MRMR), Joint Mutual Information (JMI), and Random Forest were used to specify the most important features. After attaining the significant results for features predicting 30 days and 1-year mortality, we developed a predictive model. The dataset contained class imbalance in both mortality outcomes. Firstly, under-sampling is applied as a class balancing method to eliminate 'class bias' in the training of the algorithm. Under-sampling reduces the number of observations from the majority class to make the data set balanced. Different ML algorithms were executed on the balanced dataset but the resultant accuracy of all the algorithms was very low because of the small number of records available for the training and testing models. It was then decided to perform oversampling. Oversampling increases the samples in the minority class to make the data set balanced. The data generated from oversampling have an expected amount of repeated observations and data generated from under-sampling is deprived of important information from the original data. This might lead to inaccuracies in the resulting performance. To combat these issues, on this imbalanced data hybride balancing was applied. In this case, the minority class is oversampled with replacement, and the majority class is under-sampled without replacement. After the class balancing, different ML algorithms i.e. multiple logistic regression (MLR), decision trees (DT), and SVM, were used for the prediction of 30 days and 1-year mortality.

Upon significance of various features to predict the 30 days and 1-year mortality, the accuracy, sensitivity, and specificity were compared for all the three algorithms. Each dataset was subdivided into two parts i.e. training and testing where 90% data was used for the training of the algorithms whereas 10% of the data was used for the testing of the algorithms. The patients who died within the 30 days or 1-year of referral were just 5-8% of the total data, which is why this dataset required class balancing. Once the data was balanced by hybrid-sampling, it included 924 observations in the STEMI group (448 observations for 30 days mortality which was 53 in original data set) and 1464 in the turndown group

(728 observations for 30 days and 1-year mortality which were 170 in original dataset). MLR, DT, and SVM were used on the training data set to build the predictive model for all the accepted and turndown patients. The 30 days mortality predictor variables for the STEMI group included: age, door to balloon time, call time, pain onset time, discharge on ticagrelor, ECG to balloon time, puncture site, sex, out of hour, activation. For 30 days and 1-year mortality prediction in the turndown group, the following variables were used, troponin, chest pain, referral site, age, gender, call time and out of hours referral. For the validation of the predictive model the accuracy, sensitivity, and specificity were calculated using equations 1 to 3.

$$Accuracy = \frac{TN+TP}{TN+TP+FN+FP} \qquad (1)$$

TP: True Positive, TN: True Negative, FP: False Positive, FN: False Negative

$$Specificity = \frac{TN}{TN+FP} \qquad (2)$$

$$Sensitivity = \frac{TP}{TP+FN} \qquad (3)$$

3. Results

3.1 Feature selection

3.1.1 Accepted group feature selection
Multiple feature selection techniques were applied to assess the importance of the features contributing to 30 days mortality in STEMI patients. First, correlation results didn't show any strong correlations between 30 days mortality and the features. According to MRMR, Random Forest, and logistic regression (age, DTBT, call time, discharge on ticagrelor, ECG to balloon time) (p<0.05) were statistically significant features.

3.1.2 Turndown group feature selection

Similar feature selection techniques were applied on the turndown dataset for predicting 30days and 1-year mortality in the turndown group. With the minor differences, all the techniques i.e. MRMR, JMI, random forest, and logistic regression classified the following features as significant to predict 30 days and 1-year mortality in the turndown group. The selected features included: age, troponin, chest pain, referral site, gender, call time, and out of hours.

3.2 Predicting 30 days mortality in accepted patients

A total of 980 patients were included in the original unbalanced dataset. The male dominance was evident in the dataset with 76% (n=673, p<0.001) with a mean age of 63 ± 40 years. Also 58% (515, p<0.001) of these

referrals were 'out of hours' (72% male, 28% females) with most referrals made on weekends. The 30 days mortality varied over the period of time as 4% to 10% for 2015 to 2018 respectively. In all the algorithms i.e. MLR, DT, and SVM, the independent variables including (age, DTBT, call time, pain time, discharge on Ticagrelor, ECG to balloon time, puncture site, activation) were tested to discover the effect on predicting 30 days mortality. In the accepted group, there could be true activations as well as false activations (when a patient was incorrectly urgently accepted for the Cathlab). The result of the 30 days mortality prediction using different ML techniques is shown in table 1.

For multiple logistic regression, the highest accuracy was 63% using 7 features (age, DTBT, call time, pain onset time, discharge on ticagrelor, ECG to balloon time, puncture site, sex) whereas the sensitivity and specificity were 67% and 74% respectively. Using decision tree, 7 features (age, DTBT, call time, pain time, discharge on ticagrelor, ECG to balloon time, puncture site) resulted in 81% accuracy. The sensitivity was 83% and specificity 80%. Whereas the SVM resulted in an overall accuracy 63% using the same 6 features as of MLR, however it performed the worst among all the three algorithms in terms of sensitivity and specificity (62% and 62% respectively).

Table 1. Accuracy, specificity, and sensitivity derived from multiple logistic regression, DT, and SVM to predict 30 days mortality in PPCI accepted group using a hybrid-sampled/balanced dataset.

Technique	Features	Accuracy	Sensitivity	Specificity
Multiple logistic regression	7 features	63%	67%	74%
Decision tree	8 features	81%	83%	80%
SVM	8 features	63%	62%	62%

It is evident that the decision tree outperformed the other techniques with the highest accuracy, sensitivity, and specificity rate to predict 30 days of mortality in the STEMI patients.

3.3 Predicting 30 days and 1-year mortality in turndown patients

A total of 1464 patients (68% male, P<0.001), mean age 65 ± 17 years referred for PPCI were turned down for catheterisation. Of these 'turned down' referrals, 66% (971) were out of hours. Out of all turned down patients, 117(10%) patients died within 30 days and a total of 174 (11.8%) patients died within 12 months. The result of 30 days and 1-year mortality prediction using different ML techniques is shown in table 2 and table 3. For predicting 30 days mortality in the turndown group, the highest

accuracy for MLR was 76% using 5 features (troponin, chest pain, referral site, age) with a sensitivity and specificity of 76% and 74% respectively. When the same analysis was performed using DT, 5 features (troponin, chest pain, referral site, age, gender) resulted in an accuracy of 80%, with a sensitivity of 88% and a specificity of 74%. Whereas the SVM resulted in an overall accuracy of 77% using the same 5 features as of DT with the sensitivity and specificity of is 79% and 75% respectively. Overall, DT performed the best, and MLR performed the worst among all the three algorithms.

To predict 1-year mortality in the turndown group, all three algorithms used the same number of features i.e. 4 features (troponin, chest pain, referral site, age). The highest accuracy for MLR was 73% whereas the sensitivity and specificity were 75% and 71% respectively. The same analysis with DT resulted in an accuracy of 80% with a sensitivity of 86% and a specificity of 75%. Whereas the SVM resulted in an overall accuracy of 72% with the sensitivity and specificity of 77% and 71% respectively. Overall, DT performed the best, and SVM performed the worst among all the three algorithms in terms of accuracy, sensitivity, and specificity to predict 1-year mortality in the turndown group.

Table 2 Accuracy, specificity, and sensitivity derived from multiple logistic regression, DT, and SVM to predict 30 days mortality in the PPCI turndown group when using hybrid-sampled balanced dataset

Technique	Features	Accuracy	Sensitivity	Specificity
Multiple logistic regression	4 features	76%	76%	74%
Decision Tree	5 features	80%	88%	74%
SVM	5 features	77%	79%	75%

Table 3. Accuracy, specificity, and sensitivity derived from multiple logistic regression, DT, and SVM to predict 1-year mortality in PPCI turndown group hybrid-sampled dataset.

Technique	Features	Accuracy	Sensitivity	Specificity
Multiple logistic regression	4 features	73%	75%	71%
Decision Tree	4 features	80%	86%	75%
SVM	4 features	72%	77%	71%

Overall, it is apparent that the DT outperformed the other two techniques in predicting 30 days mortality in both accepted and turndown patients.

4. Discussion

Suggesting the possible features or factors that could affect and predict the 30 days and 1-year mortality of the patients referred for the PPCI could be beneficial for clinicians. Predicting mortality for PPCI patients might help clinicians to prioritize, triage, and patient management, etc. This might also aid clinicians in decision making for accepting and turning down the patients for the CathLab. Given the importance and cruciality of PPCI picking the right patients is important.

In this research, the logistic regression, decision tree, and SVM were used to perform analysis with a PPCI patient dataset. Data imputation was performed for the data rows with 2 or more than two missing values. The amount of data imputed to each column varies from 200-400 for different columns i.e. DTBT, 30 days mortality, 1-year mortality, call time, pain time. The prediction results from the unbalanced dataset exhibited a great proportion of bias as very few observations of patients with 30 days mortality data were available. To overcome the bias issue, different class balancing techniques were used. The prediction results in terms of accuracy from the under-sampled dataset didn't exhibit a great outcome because of the small sample size. After balancing the dataset and dividing 90% data for training and 10% for testing, the ML algorithms were applied first on the accepted group, then on turndown group to predict mortality. It was observed that the oversampled balanced data performed better than under-sampled whereas the hybrid sampling performed best to predict the mortality for the patients referred to the Cath lab. Decision Tree outperformed the rest of the algorithms in predicting mortality in both accepted and turndown groups. On the other hand, SVM performed worst in almost all the cases and logistic regression remained intermediate in predicting mortality.

The overall accuracy of the DT predictive model is 83% whereas the sensitivity is 81% and specificity is 80% for the accepted group. For predicting 30 days mortality in the turndown group DT model showed an accuracy of 80% whereas the sensitivity is 88% and specificity is 74% which is the highest. Furthermore, for 1-year mortality prediction, DT again exhibits the highest accuracy (80%) with the sensitivity and specificity of 86% and 75% respectively. These results demonstrate that using different ML algorithms on PPCI referral features we can predict the 30 days and 1-year mortality of the patients referred to PPCI.

5. Conclusion

The PPCI patient records can be used to predict the 30 days and 1-year mortality of patients referred to the Cath lab by using different ML algorithms. The patient's referral feature like (age, sex, door to balloon time, ECG to balloon time, call time, pain time, activation status, chest pain) together, can be the predictor of mortality in patients. Different ML algorithms depicted different accuracy levels for the prediction. However, no ML algorithms perform well with the unbalanced and under-sampled datasets. However, with a hybrid balanced dataset, the decision tree outperformed the other algorithms to predict the mortality of PPCI referred patients. The predictor model overall accuracy for accepted and turndown group is 83% and 80% respectively.

Disclaimer

The views and opinions expressed in this paper do not necessarily reflect those of the European Commission or the Special EU Programmes Body (SEUPB).

Funding

References

[1]   K. H. Scholz et al., "Impact of treatment delay on mortality in ST-segment elevation myocardial infarction (STEMI) patients presenting with and without haemodynamic instability: results from the German prospective, multicentre FITT-STEMI trial," Eur. Heart J., vol. 39, no. 13, pp. 1065–1074, Apr. 2018.

[2]   A. Krishnamurthy et al., "Clinical outcomes following primary percutaneous coronary intervention for ST-elevation myocardial infarction according to sex and race," Eur. Hear. J. Acute Cardiovasc. Care, p. 204887261773580, 2017.

[3]   M. Motwani et al., "Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: A 5-year multicentre prospective registry analysis," Eur. Heart J., vol. 38, no. 7, pp. 500–507, 2017.

[4]   R. Shouval et al., "Machine learning for prediction of 30-day mortality after ST elevation myocardial infraction: An Acute Coronary Syndrome Israeli Survey data mining study," Int. J. Cardiol., vol. 246, pp. 7–13, Nov. 2017.

[5]   T. Song et al., "Usefulness of the heart-rate variability complex for predicting cardiac mortality after acute myocardial infarction," BMC Cardiovasc. Disord., vol. 14, May 2014.

[6]   A. J. Steele, S. C. Denaxas, A. D. Shah, H. Hemingway, and N. M. Luscombe, "Machine learning models in electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease," 2018.

Address for correspondence:
Aleeha Iftikhar
Computing Engineering and Build Environment, Ulster University Jordanstown, Northern Ireland, UK
Iftikhar-a1@ulster.ac.uk