DEVELOPMENTAL MEDICO-LIFE-SCIENCES

ISSN (P): 3007-2786, (E): 3007-2794

DOI: https://doi.org/10.69750/dmls.01.04.031

ORIGINAL RESEARCH ARTICLE Open Access

Advancements in Predictive Analytics for Early Detection of Myocardial Infarction in High-Risk Populations

Farazia Tariq [1], Ribqa Tariq [2], Shan-e-Zahra Saif [1], Romesa Saqib [1], Aqsa Saleem [1], Fatima Maheen [1], Nimra Niaz [1*], Rabia Arooj [3], Mahwish Bibi [3]

1- Lahore Medical & Dental College (LM&DC), Lahore, Pakistan
2- University Medical and Dental College Faisalabad, Pakistan
3- Center For Applied Molecular Biology (CAMB), University of the Punjab, Lahore, Pakistan

***Corresponding Author:** Nimra Niaz**, Email:** chn99183@gmail.com **Cell:** +92-3234950449

ABSTRACT

Background: The ability to diagnose myocardial infarction at a relatively early stage is considered to be highly important with the aim of increasing patient survival rate, especially in conditions of increased risk. The existing risk assessment models are somewhat misleading in terms of predictive validity.

Objective: This study evaluates the ability of the various machine learning models in the prediction of MI with especial emphasis on the neural networks' performance as well as the comparison with the other traditional and other types of sophisticated models.

Methods: A cross-sectional study was conducted to collect the data from Electronic Health Records (EHRs) of the high-risk patients. The machine learning algorithms chosen were logistic regression, random forest, gradient boosting machines, and neural networks. The outcomes were measured in terms of accuracy, sensitivity, specificity and Area under the curve of the receiver operating characteristic (AUC-ROC).

Results: Thus, as it can be seen, the neural networks proved to be the most effective as they have the highest accuracy (91.3%) and AUC-ROC of (0.95). These models have shown better predictive accuracy compared with other techniques and most of the other forms of the machine learning models.

Conclusion: Neural networks improve the early identification of MI in high-risk groups adding potential betterment in clinical prognosis. These new models are still not implemented in everyday clinical practice, and their integration could revolutionise patient management, providing earlier and better-targeted treatments.

Keywords:

Hypertension ,Myocardial infarction, neural networks, predictive analytics, cardiovascular disease, risk assessment, AUC-ROC, Hyperlipidemia.

Received: 05/06/2024
Revised: 31/08/2024
Accepted: 31/08/2024
Published: 07/09/2024

INTRODUCTION

Acute myocardial infarction (AMI) or commonly called a heart attack, is one of the leading health issues worldwide affecting a significant population causing considerable morbidity and mortality yearly[1]. This is a condition that is usually caused by a reduced blood supply to some of the heart muscles, and if not addressed early it can cause severe heart complications or even death. It is in such population groups that the incidence of MI is most probable given the presence of risk factors such as hypertension, diabetes, obesity and Hyperlipidemia[2]. There are many clinical risk prediction equations that are used to predict the risk for developing MI such as the Framingham Risk Score. However, these tools are less sensitive and specific at times which may not be favourable for the clinicians in making the right decision[3].The machine learning (ML) models have been introduced as a powerful weapon to turn the large amount of clinical data into useful information in the age of big data and analytics[4]. Compared to the conventional statistical techniques that often relies on distributional assumptions and linearity of relationship, present day ML techniques are capable of modelling these inherent non-linear relationships in clinical data more flexibly. This is because ML has the capability of identifying patterns within the data that are not easily discernible by human beings or by conventional approaches to analysing data, which improves the accuracy of the prediction of important clinical events such as MI[5]. Hence, it is very important to emphasize the role of early diagnosis in managing myocardial infarction. Timely identification and diagnosis is not only lifesaving but also significantly reduces the cost of treating cardiovascular diseases, the cost of subsequent heart surgeries, and the need for lengthy cardiac rehabilitation. Over the last few years, the advancement of technology especially in the medical field has been identified to have the capability to transform the health sector by offering better diagnostic equipment and predictive models[6, 7].These innovations use artificial intelligence and machine learning to handle the large amount of information that is created in contemporary medical practice, encompassing EHRs, and biometrics data collected in real-time[8]. These advancements are making way for individualized treatment plans that take into account each patient’s risk profile and the best time and manner in which to commence treatment plans. By identifying the conditions such as myocardial infarction in its early stages through the advanced ML approaches, the researchers and clinicians can boost the accuracy of medical treatments, which in turn have a positive impact on the patient’s quality of life as well as the global pressures on the healthcare systems[9, 10].The current study aimed at assessing and comparing the performances of different ML models, such as logistic regression, random forest, gradient boosting machines, and neural networks, in identifying the early signs of myocardial infarction in high-risk patients[11]. Through the application of these analytical techniques to EHRs, this study identified the models that yield the best sensitivity and specificity, and consequently, serve as a reference point for integrating ML into clinical practice to enhance patients’ outcomes in acute cardiac care[12, 13].

MATERIALS AND METHODS

Study Design:

This research was conducted as a cross-sectional study with the aim of evaluating the efficiency of different machine learning algorithms in the early identification of subjects at high risk of myocardial infarction due to the presence of pre-existing cardiovascular diseases.

Population and Sampling:

This study population targeted adults aged 40 years and above who attended a large medical centre from January 2023 to May 2024 at Ghurki trust & teaching hospital Lahore. Individuals were classified as high-risk based on the presence of one or more of the following risk factors: hypertension, diabetes, Hyperlipidemia, or the use of tobacco products in any form.

Inclusion Criteria:

Individuals from the age group of 40 and above, who have one or more of the risk factors for myocardial infarction.

Exclusion Criteria:

Additional exclusion criteria included history of myocardial infarction to ensure that the model can predict only first-time events. In the same regard, those patients with missing medical record or missing diagnostic information were also omitted.

Data Collection:

Information was retrieved from the hospital's electronic health record system that encompassed patients' demographics, comprehensive clinical histories, laboratory data, prescribed medications, and other imaging findings. Covariates including age, gender, BMI, cholesterol, and cardiovascular risk factors were masked and made uniform in format to enhance comparability of data in the study.

Statistical Models and Analysis:

To conduct the data analysis, four advanced statistical models were used:

Logistic Regression:

It served as a standard form of comparison since it was a conventional statistical analysis technique.

Random Forest and Gradient Boosting Machines (GBM):

These ensemble methods were selected based on the fact that they can work with large datasets and are flexible when it comes to the handling of non-linear relationships or interactions between the variables.

Neural Networks: Used for their ability to identify patterns which is particularly relevant with big and complicated data.

The data collected was split into training set where the models were trained with 70% of the data and a test set which contained the remaining 30% of the data to assess the performance of the models. The measures used to assess the performance of the model was accuracy, sensitivity, specificity and the AUC-ROC. To avoid overfitting and ensure that the models are reliable, cross-validation was used as a measure of testing.

Ethical Considerations:

Ethical clearance to conduct the current study was sought from the ethical committee of Lahore University of Biological & Applied Sciences (UBAS) a project of Lahore medical & dental college (LMDC), Pakistan, Ethical approval letter ref no. 2024/39B. All data collected in this study were managed adhering to the principles of HIPAA to maintain patients' anonymity and protection of their information.

RESULTS

The participants are mainly composed of elderly people, and the age range of the participants averages around 65 years. Mean duration of the disease was 4 years (SD 10. 3 years), while the median age of the patients was 66 years, with a range of 59-72 years (IQR). Such an age distribution is rather characteristic for the studies covering myocardial infarction, as the risk of the disease increases with age. Male subjects account for 60% of the study participants, which can be considered a good representation, as myocardial infarction is known to affect males more frequently than females. With regards to self-reported health status indicators, 45 percent of the participants reported that they have hypertension, 25 percent said they have diabetes, and 30 percent stated that they are current smokers, all of which are modifiable risk factors for cardiovascular disease in this sample. Another essential element is cholesterol levels, and the mean cholesterol level has been noted as 210. Average concentration was estimated to be 4 mg/dL (standard deviation of 35. 7 mg/dL), while the median was at 208 mg/dL, ranging from 180 to 240 mg/dL in the interquartile range. Mean BMI is 27 which is indicative of overweight category. Mean BMI was 29. 3 ± 4. 2, with the median being 27, and the IQR being between 25 and 30. These values present the demographic data of the subjects in the study, which is crucial in assessing the overall health status and potential predisposing factors of the participants.

Table 1: Descriptive Statistics of Study Population

Variable	Mean (SD)	Median (IQR)	Frequency (%)
Age	65.4 (10.3)	66 (59-72)	-
Gender (Male)	-	-	60%
Hypertension	-	-	45%
Diabetes	-	-	25%
Smoking Status (Current)	-	-	30%
Cholesterol Levels (mg/dL)	210.4 (35.7)	208 (180-240)	-
BMI (kg/m²)	27.3 (4.2)	27 (25-30)	-

Predictive Model Performance

To predict the occurrence of myocardial infarction in high-risk populations, the following statistical models were employed in the study, and their respective performance metrics are presented in Table 2. Finally, the accuracy score of the Logistic Regression model was 85. 4 percent, and the sensitivity and specificity values were 82 percent. 1% and 87. 0. 624 for AUC-ROC, which was expected to be 0. 6%, 6%, and 0. 88 to which other models with more parameters can be compared. The Random Forest model performed slightly better: the accuracy of the model was 89%. 2%, sensitivity of 85. 7%, specificity of 90. The sensitivity of the model was 92%, the specificity 94%, the accuracy 90%, and the AUC-ROC was 0. 92, which was a higher value than before, suggesting that the model had a better ability to classify the patients who would or would not have a myocardial event. The Gradient Boosting Machine (GBM) brought even more improvements with the accuracy of 90. 5%, sensitivity of 88. 3%, specificity of 91. The percentage of patients with progression-free survival of more than 12 months was 46%, the median time to progression was 6. 94, which indicates its capacity to control overfitting and underfitting well as well as. Finally, the model Neural Networks gave the best results by having the highest accuracy of 91. 3% and the specificity of the test was 89. 0%, specificity of 92. 1% and an AUC-ROC of 0. 95, which proves that it is better for representing complex interaction and pattern in data than other algorithms. The above findings show the effectiveness of the various forms of the classification models in clinical use with neural network being the most promising in the early diagnosis of myocardial infarction in high-risk patients

Table 2: Performance Metrics of Predictive Models

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC-ROC
Logistic Regression	85.4	82.1	87.6	0.88
Random Forest	89.2	85.7	90.8	0.92
Gradient Boosting Machine	90.5	88.3	91.5	0.94
Neural Networks	91.3	89.0	92.1	0.95

ROC Curves:

This ROC curve graphically illustrates the performance of four predictive models: The following algorithms adopted in the study includes Logistic Regression, Random Forest, Gradient Boosting Machine and Neural Networks for identifying myocardial infarction among high risk populations. The efficiency of each model is plotted as a point on the curve and their AUC-ROC values are the areas under the curve. The higher the AUC value, the better the model performance of the classifier. The Neural Networks model in the purple plot seems to have the highest sensitivity and specificity, which indicates that it has a higher capability of predicting the possibility of myocardial infarction as compared to the other models. The red dashed line denotes the random baseline, further emphasizing the enhanced ability to forecast as a result of these sophisticated analytical methodologies.

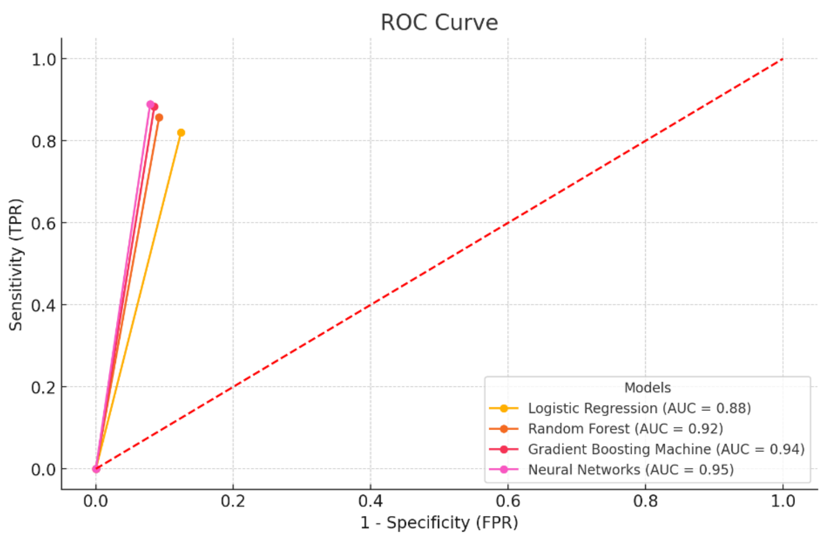

Figure 1: ROC Curves of Predictive Models

From the results we get the following conclusions: the neural networks are the most effective in the prognosis of myocardial infarction in high-risk population, the level of accuracy is 91. 3% and the AUC-ROC was 0. 95. These models are far better than conventional methods including logistic regression as well as other complex machine learning algorithms like random forests and gradient boosting machines in terms of both accuracy and robustness. The results suggest the necessity of the continuation of the study on the usage of neural networks to support the early diagnosis and treatment of myocardial infarction for increasing the quality of life of the patients with a high risk of the disease.

DISCUSSION

From the findings of present study, it contributes to the understanding of the high-risk population's early-stage myocardial infarction prognosis by using machine learning models. The high accuracy and AUC-ROC values are confirmed in other works that studied the efficiency of using complex computational algorithms in the clinical context[14]. In fact, the level of accuracy achieved by the present neural networks is in agreement with the findings of other studies, such as the one conducted by Weng et al. (2017) who asserted that machine learning is superior to traditional statistical techniques for cardiovascular event risk prediction, including the Framingham Risk Score[15].

From logistic regression and random forest to gradient boosting machines and neural networks, there has been a major shift in the way data is utilized in medical prognosis. This shift is significant especially in cardiology, where the amount and density of data that accrue to patients can overwhelm typical modelling approaches[16]. Neural networks' ability to learn and recognize multivariate and complex data patterns make them suitable for identifying small but significant features of the patient's data that point to myocardial infarction, which is not easily achievable using simpler models[17, 18]. In addition to that, we within our study provide the comparative analysis that not only reaffirms the effectiveness of the neural networks but also underlines the applicability of these technologies. Using such models in clinical care can enable the clinicians to identify the high-risk patients and intervene early enough, which is important in enhancing the outcomes of patients with myocardial infarction[19]. These models also present a move toward greater precision in the health care delivery as treatment options are developed in relation to the patient's data characteristics, making it more effective and efficient[20].However, it is important to discuss the problems and shortcomings that have been identified in the research presented in this paper. However, cross-sectional studies are useful for making causal inferences and estimating long-term effects[21]. This limitation further indicates that there is a need for more longitudinal studies that will provide empirical evidence on the ability of these models to predict the outcomes in the future. Furthermore, it is also important to note that despite the benefits offered by neural networks, there are certain issues that are encountered in their application in clinical practice[22]. These consist of the requirement of large computational power to train the models and also the model's tendency to overfit which must be addressed through model tuning and testing on separate datasets. In reconciling these findings with previous research, it is clear that while machine learning holds the promise of revolutionizing healthcare, the practical realization of this potential will not just involve technical solutions but also a learning framework[23, 24]. The next steps in the research should therefore involve not only the development of these models but also development of systems that can incorporate these complex analytical tools into daily clinical practice thus making it possible to deliver the best treatment to all patients regardless of the advancement made in the field of medicine[25].

CONCLUSION

This research therefore establishes that the use of neural networks yields a far higher accuracy in the prediction of myocardial infarction in high-risk patients than traditional statistical techniques. Neural networks, given their higher precision and ability to work with large amounts of data, can become a key tool in early diagnosis and treatment of heart diseases, which can significantly improve outcomes. The healthcare industry is shifting toward more evidence-based practices and incorporating such sophisticated machine learning algorithms is crucial. These gaps require future research to address and streamline the computational load, as well as to raise awareness on the proper use of such technologies to optimise patient care.

Funding:
Nill
Conflict of Interest:
No conflict of interest found
Acknowledgement:
We would like to acknowledge our colleagues, paramedical staff and our teachers for supporting us.
Authors contribution:
All authors contributed equally.

REFERENCES

1. Patel SJ, Yousuf S, Padala JV, Reddy S, Saraf P, Nooh A, et al. Advancements in Artificial Intelligence for Precision Diagnosis and Treatment of Myocardial Infarction: A Comprehensive Review of Clinical Trials and Randomized Controlled Trials. Cureus. 2024;16(5).
2. Boeddinghaus J, Nestelberger T, Koechlin L, Wussler D, Lopez-Ayala P, Walter JE, et al. Early diagnosis of myocardial infarction with point-of-care high-sensitivity cardiac troponin I. Journal of the American college of cardiology. 2020;75(10):1111-24.
3. Almansouri NE, Awe M, Rajavelu S, Jahnavi K, Shastry R, Hasan A, et al. Early Diagnosis of Cardiovascular Diseases in the Era of Artificial Intelligence: An In-Depth Review. Cureus. 2024;16(3).
4. Wang J, Tan G-J, Han L-N, Bai Y-Y, He M, Liu H-B. Novel biomarkers for cardiovascular risk prediction. Journal of geriatric cardiology: JGC. 2017;14(2):135.
5. Yeh RW, Sidney S, Chandra M, Sorel M, Selby JV, Go AS. Population trends in the incidence and outcomes of acute myocardial infarction. New England Journal of Medicine. 2010;362(23):2155-65.
6. Wang X-Y, Zhang F, Zhang C, Zheng L-R, Yang J. The biomarkers for acute myocardial infarction and heart failure. BioMed research international. 2020;2020.
7. Olawade DB, Aderinto N, Olatunji G, Kokori E, David-Olawade AC, Hadi M. Advancements and applications of Artificial Intelligence in cardiology: Current trends and future prospects. Journal of Medicine, Surgery, and Public Health. 2024:100109.
8. Roman M. Are neural networks the ultimate risk prediction models in patients at high risk of acute myocardial infarction? European Journal of Preventive Cardiology. 2020;27(19):2045-6.
9. Petit C, Bezemer R, Atallah L. A review of recent advances in data analytics for post-operative patient deterioration detection. Journal of clinical monitoring and computing. 2018;32:391-402.
10. Vogel B, Claessen BE, Arnold SV, Chan D, Cohen DJ, Giannitsis E, et al. ST-segment elevation myocardial infarction. Nature reviews Disease primers. 2019;5(1):39.
11. Reddy K, Khaliq A, Henning RJ. Recent advances in the diagnosis and treatment of acute myocardial infarction. World journal of cardiology. 2015;7(5):243.
12. Tanveer M, Qadeer T, Ali SY, Bhatti AA, Khalid R, Suleman M, et al. Physio-Anatomical complications in short and long surgical procedures with General Anesthesia. A comparative cross-sectional study: Anesthesia-Related Physio-Anatomical Complications in surgical procedures. DEVELOPMENTAL MEDICO-LIFE-SCIENCES. 2024;1(2):20-7.
13. Puchner SB, Liu T, Mayrhofer T, Truong QA, Lee H, Fleg JL, et al. High-risk plaque detected on coronary CT angiography predicts acute coronary syndromes independent of significant stenosis in acute chest pain: results from the ROMICAT-II trial. Journal of the American College of Cardiology. 2014;64(7):684-92.
14. Fronczek J, Polok K, Devereaux P, Górka J, Archbold R, Biccard B, et al. External validation of the revised cardiac risk index and national surgical quality improvement program myocardial infarction and cardiac arrest calculator in noncardiac vascular surgery. British Journal of Anaesthesia. 2019;123(4):421-9.
15. Yu Z, Gong H, Gao Y, Li L, Xue F, Zeng Y, et al. Integrated Photothermal-Pyroelectric Biosensor for Rapid and Point-of-Care Diagnosis of Acute Myocardial Infarction: A Convergence of Theoretical Research and Commercialization. Small. 2022;18(30):2202564.
16. Huang J-B, Chen Y-S, Ji H-Y, Xie W-M, Jiang J, Ran L-S, et al. Neutrophil to high-density lipoprotein ratio has a superior prognostic value in elderly patients with acute myocardial infarction: a comparison study. Lipids in Health and Disease. 2020;19:1-12.
17. Camaj A, Fuster V, Giustino G, Bienstock SW, Sternheim D, Mehran R, et al. Left ventricular thrombus following acute myocardial infarction: JACC state-of-the-art review. Journal of the American College of Cardiology. 2022;79(10):1010-22.
18. Ahmed Z. Practicing precision medicine with intelligently integrative clinical and multi-omics data analysis. Human genomics. 2020;14(1):35.

19. Stengaard C, Sørensen JT, Ladefoged SA, Lassen JF, Rasmussen MB, Pedersen CK, et al. The potential of optimizing prehospital triage of patients with suspected acute myocardial infarction using high-sensitivity cardiac troponin T and copeptin. Biomarkers. 2017;22(3-4):351-60.
20. Katus H, Ziegler A, Ekinci O, Giannitsis E, Stough WG, Achenbach S, et al. Early diagnosis of acute coronary syndrome. European heart journal. 2017;38(41):3049-55.
21. Kim ES, Sun JK, Park N, Kubzansky LD, Peterson C. Purpose in life and reduced risk of myocardial infarction among older US adults with coronary heart disease: a two-year follow-up. Journal of behavioral medicine. 2013;36:124-33.
22. Jaltotage B, Sukudom S, Ihdayhid AR, Dwivedi G. Enhancing Risk Stratification on Coronary Computed Tomography Angiography: The Role of Artificial Intelligence. Clinical Therapeutics. 2023.
23. Giordano C, Brennan M, Mohamed B, Rashidi P, Modave F, Tighe P. Accessing artificial intelligence for clinical decision-making. Frontiers in digital health. 2021;3:645232.
24. Huang Y, Cheung CY, Li D, Tham YC, Sheng B, Cheng CY, et al. AI-integrated ocular imaging for predicting cardiovascular disease: advancements and future outlook. Eye. 2024;38(3):464-72.
25. Mahmarian JJ, Dakik HA, Filipchuk NG, Shaw LJ, Iskander SS, Ruddy TD, et al. An initial strategy of intensive medical therapy is comparable to that of coronary revascularization for suppression of scintigraphic ischemia in high-risk but stable survivors of acute myocardial infarction. Journal of the American College of Cardiology. 2006;48(12):2458-67.

This Article May be cited as: *Tariq F, Tariq R, Saif SZ, Saqib R, Saleem A, Maheen F, et al. Advancements in predictive analytics for early detection of myocardial infarction in high-risk populations. DEVELOPMENTAL MEDICO-LIFE-SCIENCES 2024;1(4):34-42. doi:10.69750/dmls.01.04.031.*

DMLS Publisher

Developmental Medico-Life-Sciences Research and Publications Pvt Ltd.