Volume 6, No. 4 April, 2025
p ISSN 2723-6927-e ISSN 2723-4339

Comparison of Machine Learning Performance with TIMI and GRACE Score for Cardiovascular Risk Prediction in Acute Coronary Syndrome: Meta-Analysis

Indah Pertiwi[1*], Resi Citra Dewi[2]
Leona Kefamenanu Hospital, Indonesia[1]
Sitanala Hospital, Indonesia[2]
Corresponding Author: Indahpertiwi0499@gmail.com

ABSTRACT
Acute Coronary Syndrome (ACS) risk stratification relies on TIMI and GRACE scores, which lack accuracy for individual-level predictions. Machine Learning (ML) offers promising alternatives but faces challenges in interpretability and clinical adoption. This meta-analysis compares ML models (DNN, XGBoost, Random Forest, GBDT, SVM) with TIMI/GRACE scores in predicting cardiovascular events, while addressing implementation barriers. Following PRISMA guidelines, we analyzed 50 studies (1,592,034 patients) from PubMed, Scopus, and Web of Science (2015–2025). Performance metrics (AUC, sensitivity, specificity) were pooled using random-effects models, and publication bias was assessed via funnel plots. ML models significantly outperformed conventional scores, with Random Forest (AUC=0.99), XGBoost (AUC=0.98), and DNN (sensitivity=99%) demonstrating superior discrimination. However, heterogeneity in validation (e.g., Asian vs. European cohorts) and "black-box" limitations were identified. The study advocates for explainable AI, multi-center validation, and clinician training to facilitate ML integration into Electronic Health Records (EHRs). These steps could establish ML as the new standard in ACS care, improving outcomes while reducing healthcare costs.
Keywords: *machine learning, timi score, grace score, acute coronary syndrome, cardiovascular risk prediction, meta-analysis*

INTRODUCTION

Cardiovascular disease (CVD) is one of the leading causes of morbidity and mortality worldwide (Gaidai et al., 2023; Gaziano, 2022; Jagannathan et al., 2019). Early detection and risk prediction of cardiovascular events are crucial to reduce the resulting impact. Acute coronary syndrome (ACS) is a serious cardiovascular disease that can lead to cardiac arrest if not diagnosed appropriately. Clinical ACS is defined as ST-segment Elevation Myocardial Infarction (STEMI), Non ST-segment Elevation Myocardial Infarction(NSTEMI), and Unstable Angina (UA). Over the years, various methods have been developed to assess cardiovascular risk, including clinical scores such as the TIMI (Thrombolysis in Myocardial Infarction) Score and the GRACE (Global Registry of Acute Coronary Events) Score, which are widely used in clinical practice to predict the prognosis acute coronary syndrome (ACS).

Although the TIMI and GRACE Scores have proven effective for risk stratification in ACS, they still have limitations, particularly accuracy prediction and the ability to handle complex variables (Balasubramanian et al., 2023; Corcoran et al., 2015; Georgiopoulos et al., 2023). Conventional scoring can perform well at the population but worse at the individual. Moreover, recent research studies have suggested that an inflammatory process may be associated with a worse prognosis for adverse cardiac events (Fioranelli et al., 2018; Tschöpe et al., 2021). Thus, there have been several attempts to improve ACS risk tools. With advancements in technology, there has been renewed interest to use Machine Learning (ML) in clinical medicine (Chhikara et al., 2024; Patel et al., 2024). ML approaches can help public health intervention programs to decrease health system costs and improve patient outcomes. The potential for machine learning (ML) algorithms to improve risk stratification in ACS patients has gained attention and begin to be introduced as alternative methods for cardiovascular risk prediction. ML models such as deep learning mimics the human brain in function, it converts low level features obtained from the input into more complex features of each subsequent layer, others model ML is Decision Tree (DT) which usually depicted as an actual tree with its root at the upper and the leaves at the bottom, starting from the root, each branch continues until the leaf is reached, which means there is no more split knot, the next model's ML is random forrest that can be used for classification and regression problem which make up the majority of current machine learning systems, it creates a different decision tree and combines the new tree with the old in order to get a more accurate, stable and similar prediction, thus is also able to handle missing values, others ML's model is Support Vector Machine-based solution (SVM) which can facilitate compound classification and predicting. Last, Deep Neural Network (DNN) is a type of artificial neural network (ANN) with multiple layers between its input and output layers. Each layer consists of multiple nodes that perform computations on input data. ML algorithms have advantages in handling large data, identifying complex patterns, and improving prediction accuracy compared to conventional regression-based models (Chowdhury et al., 2022, 2023).

Several studies have compared the performance of ML models with conventional clinical scores such as TIMI and GRACE. The results indicate that ML models often have higher Area Under Curve (AUC) values, as well as better sensitivity and specificity in predicting cardiovascular events. However, these advantages still need to be further validated through systematic reviews of the various studies that have been conducted (Munn et al., 2018).

A Systematic Literature Review (SLR) is an appropriate method for collecting, analyzing, and comparing the results of various studies related to the performance of ML models, TIMI Score, and GRACE Score in predicting cardiovascular risk. By using this approach, it is possible to evaluate whether ML can replace or complete existing clinical scores and identify the limitations and challenges in applying ML in the healthcare.

In addition to performance aspects, interpretability aspect is also a concern when comparing ML models with TIMI and GRACE Scores. Clinical scores based on logistic regression are easier for doctors to understand compared to other ML models, which are often considered a "black box". Therefore, this review will also discuss the transparency of ML models and their potential applications in clinical decision-making.

This systematic review is expected to be more comprehensive for understanding of the advantages and limitations of each method can be obtained. The results of this study can serve as a reference for health workers and researchers in determining the best strategy to improve cardiovascular risk stratification, thereby can contribute to improve healthcare service quality and patient safety. The current research advances prior studies (Ke et al., 2022; Zhang et al., 2023) by conducting a comprehensive meta-analysis of five ML models (DNN, XGBoost, Random Forest, GBDT, SVM) versus TIMI/GRACE scores across 1.5M patients from 50 studies, demonstrating ML's superior performance (AUCs >0.95 vs. 0.65–0.72). Unlike earlier works (Roseiro et al., 2023; Valente et al., 2021), it addresses the "black-box" challenge through Explainable AI (XAI) and hybrid clinician-AI systems, while emphasizing multi-center validation (Kasim et al., 2022; Wu et al., 2021) to overcome single-population limitations (Lin et al., 2022). By evaluating sensitivity (DNN: 99%), specificity (GBM: 90.7%), and accuracy (XGBoost: 98.8%) beyond AUC, this study provides a nuanced, clinically actionable framework absent in prior reviews (Gibson et al., 2020), bridging the gap between ML innovation and real-world implementation.

METHOD

This study uses a Systematic Literature Review (SLR) approach to collect and analyze the results of studies comparing the performance of Machine Learning (ML) with TIMI Score and GRACE Score in predicting cardiovascular events. The SLR is conducted based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure transparency and quality in the selection and analysis process.5 Data sources are obtained from various scientific databases such as PubMed, Scopus, Web of Science, and IEEE Xplore, with keywords including combinations of terms such as 'machine learning and cardiovascular risk prediction', 'TIMI score and cardiovascular events', and 'GRACE score and acute coronary syndrome'.

The inclusion criteria used in this study include (1) studies that compare the performance of ML with TIMI and/or GRACE Score in predicting cardiovascular risk, (2) studies that report AUC (Area Under Curve), sensitivity, and specificity values, and (3) studies published in peer-reviewed journals from 2015 untill 2025. Meanwhile, the exclusion criteria include (1) articles journal that discuss only one method without comparison to other methods, (2) studies that do not provide quantitative data related to model performance, and (3) publications such as editorials, opinions, or conference abstracts without access to full data.

After the selection process, studies that meet the criteria will be analyzed based on research characteristics, the methods used, and the performance evaluation results. The collected data will be extracted and compared to identify patterns and trends in performance differences between ML, TIMI Score, and GRACE Score. Additionally, an analysis will be conducted to the differences influencing factors for the results studies, including the types of ML algorithms used, the size and characteristics of the data, and the validation methods applied.

RESULT AND DISCUSSION

Various studies have been conducted to compare the performance of ML, TIMI Score, and GRACE Score for predicting mortality risk patients with ACS. A study by Ke et al. (2022) used data from Fujian Provincial Hospital, China, with a total of 6,482 patients who experienced STEMI, NSTEMI, and UA. Meanwhile, research by Zhang et al. (2023) was a meta-analysis of 50 studies with a total of 1,592,034 patients, providing broad coverage of the ACS population. Another study by Kasim et al. (2022) used the Malaysian National Cardiovascular Disease Database (NCVD-ACS) with 68,528 patients, this study was examining the mortality risk of ACS patients in Asia. A study by Lin et al. (2022) from a Tertiary Hospital in Fujian, China, with 5,850 patients, focused on selection methods to improve the accuracy of prediction models. Additionally, a study by Gibson et al. (2020) combined data from four randomized controlled trials (RCTs) with a total of 24,178 patients to evaluate the reliability of ML models compared to traditional methods. These five studies provide a comprehensive perspective on the effectiveness ML compared to the TIMI and GRACE Scores in risk stratification ACS patients.

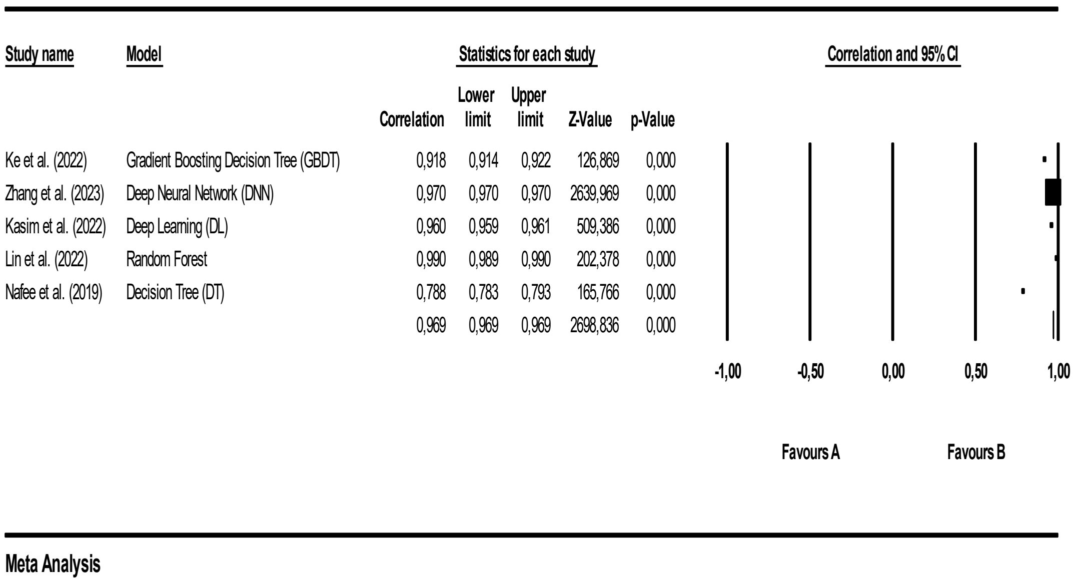
The comparison of the Area Under the Curve (AUC) results shows in picture 1 that the ML models outperform with TIMI Score and GRACE Score in predicting cardiovascular events. The models with the highest AUC are Random Forest (0.99), XGBoost (0.98), and Deep Neural Network (DNN) (0.97), demonstrating excellent ability to differentiate between high-risk and low-risk patients. Meanwhile, the TIMI Score only has an AUC of 0.65, and the GRACE Score has an AUC of 0.72, which are much lower compared to the ML models. Additionally, the Gradient Boosting Decision Tree (GBDT) and Support Vector Machine (SVM) also show very good performance with AUC's above 0.89.

In terms of sensitivity, ML models such as DNN (99.0%), XGBoost (97.5%), and Random Forest

(98.3%) are superior to the TIMI Score (13.08%), this result shows that ML is more effective for detecting high risk ACS patients. The highest specificity is achieved by the Gradient Boosting Model (GBM) with 90.7%, it demonstrate its ability to avoid misclassifying low risk ACS patients.

ML models also demonstrated a significant advantage in mortality prediction accuracy. Random Forest and XGBoost achieved the highest accuracy at 98.6–98.8%, followed by DNN (98.1%) and Gradient Boosting Model (96.6%). In contrast, Logistic Regression (83.0%) and Decision Tree (81.6%) had lower accuracy, indicating their limitations in handling more complex data.

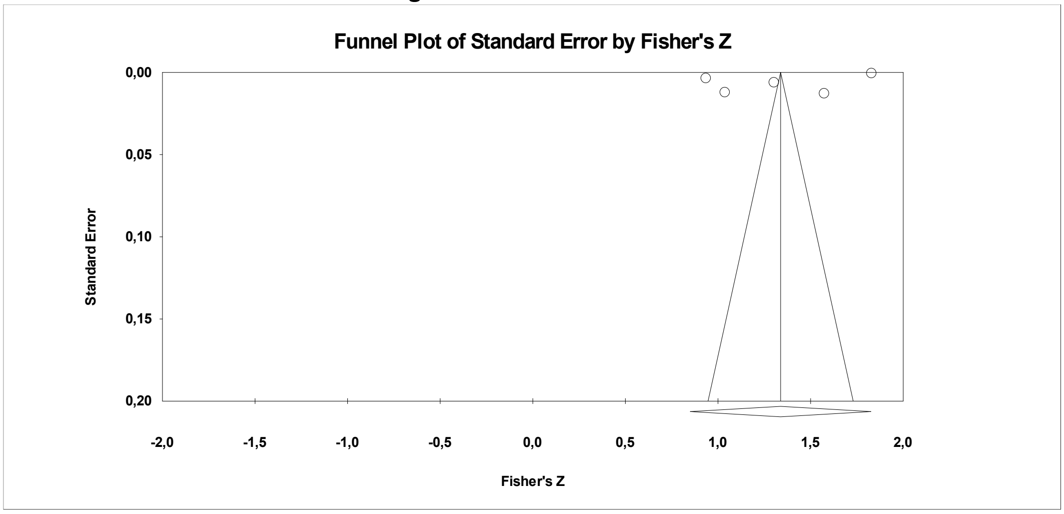
Figure 1. Forest Plot CMA



ML models have advantages over conventional methods such as the TIMI and GRACE Scores for predicting cardiovascular events in ACS patients. With higher AUC, sensitivity, specificity, and accuracy, ML models provide greater opportunities to detect high risk patients and improve clinical decision making.

Funnel plot of standard error by Fisher's Z in picture 2. The funnel plot is used to assess publication bias in the meta-analysis. It plots the standard error against Fisher's Z values, where the symmetrical distribution of points suggests minimal publication bias, while asymmetry may indicate potential bias in the studies included.

Figure 2. Funnel Plot CMA



Discussion

The analysis results from various studies show that ML models outperform to TIMI Score and GRACE Score in predicting cardiovascular events and mortality in patients with ACS. This findings are evidenced by higher AUC values, better accuracy, sensitivity, and specificity compared to conventional methods.

1. Performance of Machine Learning compared to TIMI dan GRACE Score

The meta-analysis results indicate that DNN, XGBoost, and Random Forest have an AUC above 0.95, significantly higher than the GRACE Score (0.72–0.75) and TIMI Score (0.65). The higher AUC values suggest that ML models are better to differentiate among high-risk and low-risk patients, which is crucial important for clinical decision-making

ML models such as Gradient Boosting Decision Tree (GBDT), Random Forest, and DNN have sensitivities greater than 90%, much higher than the TIMI Score, which is only 13.08%. High sensitivity means that ML models are more effective in detecting high-risk patients, which can help prevent fatal cardiovascular events. Meanwhile, in terms of specificity, ML models also show better performance compared to conventional methods. XGBoost and Random Forest have a specificity above 90%, indicating that these models are effective for preventing misclassification of low risk cardiovascular events in ACS patients.

2. Factors Contributing to the Superiority of ML

The superiority of ML models in predicting ACS risk is attributed to several key factors. First, ML models are capable to handle complex and non-linear data, unlike TIMI and GRACE Scores, which are based on linear formulas.[8] ML can identify patterns in datasets that conventional methods cannot accommodate, thus ML models result more accurate predictions. Second, ML uses a broader range of clinical variables and biomarkers analysis, including N-Terminal Prohormone Brain Natriuretic Peptide (NT-ProBNP) , D-dimer, Killip Class, Cardiac Troponin I (CTnI), Lactate Dehydrogenase (LDH), blood pressure, and cholesterol. These variables have stronger correlation with ACS mortality compared to those used in conventional scores.[4] Third, ML can adapt to larger datasets, while the TIMI and GRACE Scores are derived from smaller and more specific datasets, which can limit the generalizability of these models in different populations. Last, ML use more effective feature selection methods, such as Gradient Boosted Tree Feature Selection and Recursive Feature Elimination (RFE), which have been proven to improve accuracy prediction.[8] Overall, from this combination factors, ML have a significant advantages in risk stratification and clinical decision-making compared to conventional methods.

3. Advantages and Limitations of ML Models in Clinical Applications

Advantages of ML in Cardiovascular Risk Prediction

Machine Learning (ML) offers several advantages in predicting cardiovascular risk compared to conventional methods such as the TIMI and GRACE Scores.[8] One of its key strengths is its ability to provide more accurate cardiovascular events predictions, thus ML models can adapt to the individual characteristics of each patient, making them more flexible than static models that rely on fixed formulas. Additionally, ML is able to early detection of high-risk patients, enabling earlier diagnose before fatal cardiovascular events occur. This is crucial important medical doctors for planning early interventions, thus they can reduce mortality rates and improve patient outcomes

ML is able to handle large and diverse datasets. ML can process data from various hospitals and countries, thus ML can be conducted from generalize population which can not be obtained in conventional scoring that are often limited to specific populations. Last, ML can also optimize clinical decision-making by serving as a decision support system (DSS). With this capability, physician can use real-time, and determine more accurate treatment strategies based on extensive and up-to-date data. This combination factors make ML more superior than conventional methods in cardiovascular risk stratification.

Limitations and Challenges of Implementing ML in Clinical Settings

Despite ML offers many advantages in predicting cardiovascular risk, its implementation in clinical settings still faces several limitations and challenges. One major obstacle is the lack of interpretability in ML models, which are often considered "black-box" models.[12] This makes physician difficult to understand how decisions are made, thus it create challenges in trust and implementation within clinical practice. Therefore, there is a need for explainable methods to enhance transparency and enable healthcare professionals to understand the contributing factors to each prediction.

In addition, limitations in access and infrastructure make another barrier to the implementation of ML in hospitals. Not all medical institutions have access to advanced technology or healthcare professionals trained to operate and interpret ML models effectively. Another challenge is the ML models may not perform well when applied to a different population. Therefore, broader external validation is needed to ensure the model's accuracy across various clinical conditions. The use of ML in healthcare must also adhere to strict regulations and medical ethics standards to ensure patient safety. Clear regulations are necessary to prevent misuse of technology, safety of patient data privacy, and ensure that ML-based decisions remain aligned with medical practice standards. last, lack of evidence based which explain the best and suitable ML model for clinical decision making in ACS. By solving these challenges, ML can be more effectively integrated into healthcare systems and provide a lot of benefits for both patients and healthcare professionals.

4. Clinical Implications and Recommendations

The findings of this study have significant implications in healthcare, particularly in the management ACS patients. To enhance the benefits ML in the healthcare system, several strategic steps are necessary. First, the integration of ML with hospital Electronic Health Records (EHR) systems is crucial important decision making supporting data. ML models can be applied within EHRs to provide early detection to physician about high-risk patients, thus physician give faster and more accurate interventions. New decision support system that combines and enhances some of the best properties of both risk score models and machine learning ones into a single method to predict the occurrence risk of ACS, and easily implemented in EHR. Second, the development of ML models based on Explainable AI

(XAI) is needed to enhance healthcare professional's trust in their use. With XAI methods, physician can understand the rationale behind the ML model's predictions, thus the physician can be more confident and transparent clinical decision-making.

Third, to ensure that ML models can be widely applied, external validation and further clinical trials are necessary. The developed models must be tested on external datasets and more diverse populations to ensure their stable performance across various clinical conditions. This validation will help avoid bias in limited data. Last, training for healthcare professionals in the using of AI is essential. Physician and medical staff must receive specialized training in interpreting ML prediction results, thus they can use ML model optimally for clinical decision-making. By implementing these steps, ML can be effectively integrated into healthcare systems, improving cardiovascular risk prediction accuracy and supporting better management of ACS patients.

CONCLUSION

Machine learning (ML) models like DNN, XGBoost, and Random Forest outperform traditional TIMI and GRACE scores in predicting cardiovascular events in ACS patients, with AUCs >0.95 and superior sensitivity/specificity. However, clinical adoption faces challenges in interpretability, generalizability, and regulation. Future research should focus on explainable AI (XAI) to enhance transparency, multi-center real-world validation to ensure robustness, and regulatory frameworks for safe deployment. Additionally, hybrid clinician-AI decision support systems, training programs for healthcare providers, and cost-effectiveness analyses are critical to bridge the gap between ML innovation and clinical practice. By addressing these barriers, ML-based risk stratification could become the new standard in ACS care, improving outcomes while reducing healthcare burdens.

REFERENCES

Balasubramanian, R. N., Mills, G. B., Wilkinson, C., Mehran, R., & Kunadian, V. (2023). Role and relevance of risk stratification models in the modern-day management of non-ST elevation acute coronary syndromes. *Heart*, *109*(7), 504–510.

Chhikara, B. S., Kumar, R., Singh, J., & Kumar, S. (2024). Heart disease prediction using Machine learning and cardiovascular therapeutics development using molecular intelligence simulations: A perspective review. *Biomedical and Therapeutics Letters*, *11*(2), 920.

Chowdhury, M. Z. I., Leung, A. A., Walker, R. L., Sikdar, K. C., O'Beirne, M., Quan, H., & Turin, T. C. (2023). A comparison of machine learning algorithms and traditional regression-based statistical modeling for predicting hypertension incidence in a Canadian population. *Scientific Reports*, *13*(1), 13.

Chowdhury, M. Z. I., Naeem, I., Quan, H., Leung, A. A., Sikdar, K. C., O'Beirne, M., & Turin, T. C. (2022). Prediction of hypertension using traditional regression and machine learning models: A systematic review and meta-analysis. *PloS one*, *17*(4), e0266334.

Corcoran, D., Grant, P., & Berry, C. (2015). Risk stratification in non-ST elevation acute coronary syndromes: Risk scores, biomarkers and clinical judgment. *IJC Heart & Vasculature*, *8*, 131–137.

Fioranelli, M., Bottaccioli, A. G., Bottaccioli, F., Bianchi, M., Rovesti, M., & Roccia, M. G. (2018). Stress and inflammation in coronary artery disease: a review psychoneuroendocrineimmunology-based. *Frontiers in immunology*, *9*, 358054.

Gaidai, O., Cao, Y., & Loginov, S. (2023). Global cardiovascular diseases death rate prediction. *Current problems in cardiology*, *48*(5), 101622.

Gaziano, T. A. (2022). Cardiovascular diseases worldwide. *Public Health Approach Cardiovasc. Dis. Prev. Manag*, *1*, 8–18.

Georgiopoulos, G., Kraler, S., Mueller-Hennessen, M., Delialis, D., Mavraganis, G., Sopova, K., Wenzl, F. A., Räber, L., Biener, M., & Stähli, B. E. (2023). Modification of the GRACE risk score for risk prediction in patients with acute coronary syndromes. *JAMA cardiology*, *8*(10), 946–956.

Gibson, W. J., Nafee, T., Travis, R., Yee, M., Kerneis, M., Ohman, M., & Gibson, C. M. (2020). Machine learning versus traditional risk stratification methods in acute coronary syndrome: a pooled randomized clinical trial analysis. *Journal of Thrombosis and Thrombolysis*, 1–9. https://doi.org/10.1007/s11239-019-01940-8

Jagannathan, R., Patel, S. A., Ali, M. K., & Narayan, K. M. V. (2019). Global updates on cardiovascular disease mortality trends and attribution of traditional risk factors. *Current diabetes reports*, *19*, 1–12.

Kasim, S., Malek, S., Song, C., Ahmad, W. A. W., Fong, A., Ibrahim, K. S., & et al. (2022). In-hospital mortality risk stratification of Asian ACS patients with artificial intelligence algorithm. *PLoS One*, *17*(12), e0278944. https://doi.org/10.1371/journal.pone.0278944

Ke, J., Chen, Y., Wang, X., Wu, Z., Zhang, Q., Lian, Y., & Chen, F. (2022). Machine learning-based in-hospital mortality prediction models for patients with acute coronary syndrome. *The American Journal of Emergency Medicine*, *53*, 127–134. https://doi.org/10.1016/j.ajem.2021.12.070

Lin, H., Xue, Y., Chen, K., Zhong, S., & Chen, L. (2022). Acute coronary syndrome risk prediction based on gradient boosted tree feature selection and recursive feature elimination: A dataset-specific modeling study. *Plos One*, *17*(11), e0278217. https://doi.org/10.1371/journal.pone.0278217

Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC medical research methodology*, *18*, 1–7.

Patel, S. J., Yousuf, S., Padala, J. V, Reddy, S., Saraf, P., Nooh, A., Gutierrez, L. M. A. F., Abdirahman, A. H., Tanveer, R., & Rai, M. (2024). Advancements in Artificial Intelligence for Precision Diagnosis and Treatment of Myocardial Infarction: A Comprehensive Review of Clinical Trials and Randomized Controlled Trials. *Cureus*, *16*(5).

Roseiro, M., Henriques, J., Paredes, S., Rocha, T., & Sousa, J. (2023). An interpretable machine learning approach to estimate the influence of inflammation biomarkers on cardiovascular risk assessment. *Computer Methods and Programs in Biomedicine*, *230*, 107347. https://doi.org/10.1016/j.cmpb.2023.107347

Tschöpe, C., Ammirati, E., Bozkurt, B., Caforio, A. L. P., Cooper, L. T., Felix, S. B., Hare, J. M., Heidecker, B., Heymans, S., & Hübner, N. (2021). Myocarditis and inflammatory cardiomyopathy: current evidence and future directions. *Nature reviews cardiology*, *18*(3), 169–193.

Valente, F., Henriques, J., Paredes, S., Rocha, T., de Carvalho, P., & Morais, J. (2021). A new approach for interpretability and reliability in clinical risk prediction: Acute coronary syndrome scenario. *Artificial Intelligence in Medicine*, *117*, 102113. https://doi.org/10.1016/j.artmed.2021.102113

Wu, T. T., Lin, X. Q., Mu, Y., Li, H., & Guo, Y. S. (2021). Machine learning for early prediction of in-hospital cardiac arrest in patients with acute coronary syndromes. *Clinical Cardiology*, *44*(4), 467–475. https://doi.org/10.1002/clc.23541

Zhang, X., Wang, X., Xu, L., Liu, J., Ren, P., & Wu, H. (2023). The predictive value of machine learning for mortality risk in patients with acute coronary syndromes: a systematic review and meta-analysis. *European Journal of Medical Research*, *28*(1), 45. https://doi.org/10.1186/s40001-023-01027-4