American Journal of Medicine and Medical Sciences 2025, 15(10): 3493-3495
DOI: 10.5923/j.ajmms.20251510.45

Neural Networks in the Prediction of Cardiovascular Disease: A Data-Driven Approach

Abdulhamid Haydarov[1,*], Nurmukhammad Mukhitdinov[2], Jamshid Jabborov[3], Dilshoda Sayfullayeva[4], Khumora Shodmonova[5]

[1]Faculty of Pediatrics, Andijan State Medical Institute, Andijan, Uzbekistan
[2]Faculty of Pharmacy, Andijan State Medical Institute, Andijan, Uzbekistan
[3]Faculty of Computer Engineering, Tashkent University of Information Technology Qarshi Branch, Uzbekistan
[4]Faculty of Pediatrics, Tashkent State Medical University, Tashkent, Uzbekistan
[5]Faculty of General Medicine, Bukhara State Medical Institute, Bukhara, Uzbekistan

Abstract Cardiovascular diseases (CVDs) remain the leading cause of morbidity and mortality worldwide, underscoring the need for reliable predictive models to enable timely preventive interventions. Traditional models such as Framingham and QRISK, while widely used, are limited in their ability to capture complex nonlinear relationships among diverse risk factors [2,5]. This study aimed to develop and validate a neural network–based predictive model for cardiovascular disease risk, incorporating demographic, clinical, and lifestyle-related variables. We retrospectively analyzed a data-set of 1,500 patients drawn from multi-institutional electronic health records, applying rigorous preprocessing techniques including normalization, feature engineering, and missing data imputation. The neural network architecture consisted of three hidden layers with ReLU activation, optimized using Adam, and trained on 70% of the dataset with validation on 15% and external testing on 15%. Predictive performance was compared with logistic regression, Framingham, and QRISK scores. The neural network achieved superior discrimination (AUROC 0.89) compared with logistic regression (0.78), Framingham (0.74), and QRISK (0.77). Calibration was excellent across deciles of risk, and decision curve analysis showed a net benefit of up to 8 avoided unnecessary interventions per 100 patients compared with Framingham, QRISK, and logistic regression models. SHAP interpretability analysis identified blood pressure, cholesterol, physical activity, family history, and stress as the most influential predictors. These findings highlight the transformative potential of neural networks in cardiovascular risk prediction, offering enhanced accuracy and supporting the paradigm shift toward precision medicine. Integration of such models into clinical workflows could significantly improve preventive cardiology practices and population health outcomes [3,10,13].

Keywords Cardiovascular Disease Prediction, Neural Networks, Machine Learning in Cardiology, Risk Stratification, Precision Medicine, Explainable Artificial Intelligence (XAI)

1. Introduction

Cardiovascular diseases (CVDs) account for nearly one-third of all global deaths, presenting a pressing public health challenge [1]. Traditional risk assessment tools such as the Framingham Risk Score and QRISK have been instrumental in guiding preventive care, yet they rely on linear assumptions and a limited set of predictors [2,5]. These constraints often lead to underestimation or overestimation of risk, particularly in diverse populations [8].

* Corresponding author:
a_xaydarov@icloud.com (Abdulhamid Haydarov)
Received: Sep. 28, 2025; Accepted: Oct. 22, 2025; Published: Oct. 25, 2025
Published online at http://journal.sapub.org/ajmms

Recent advances in machine learning, particularly neural networks, provide an opportunity to overcome these limitations. By modeling complex, nonlinear interactions among diverse risk factors, they enable more accurate predictions [3,7,16]. Neural networks have demonstrated superior predictive performance across multiple domains of medicine, including oncology, neurology, and cardiology [4,9]. However, questions remain about their clinical applicability, interpretability, and external generalizability.

This study investigates the utility of a neural network–based approach for predicting CVD risk. By integrating demographic, clinical, and lifestyle-related variables, we aimed to assess whether neural networks could outperform conventional models, evaluate their interpretability, and explore their potential role in advancing precision cardiology [6,11].

2. Materials and Methods

Study Design and Population

This was a retrospective, multi-center study analyzing anonymized electronic health records from 1,500 patients aged 30–80 years, collected between 2015 and 2023 from three tertiary care hospitals. Patients with a prior history of cardiovascular disease at baseline were excluded to ensure incident outcome prediction. The external test set was created as a temporally separated cohort from the latest year of data across the three hospitals, ensuring independence from the training/validation sets.

Data Collection and Variables

Extracted variables included:

- **Demographics:** age, sex, ethnicity
- **Clinical parameters:** systolic and diastolic blood pressure, body mass index, fasting glucose, lipid profile
- **Lifestyle factors:** smoking, alcohol use, diet quality, physical activity level, sleep duration, stress indicators
- **Family history:** first-degree relatives with CVD

Outcomes were defined as first incident myocardial infarction, stroke, or cardiovascular death, confirmed by ICD-10 coding and discharge summaries [2,8].

Data Preprocessing

- Missing data were imputed using k-nearest neighbors.
- Continuous variables were normalized (z-scores).
- Categorical variables were one-hot encoded.
- Feature engineering incorporated composite lifestyle scores (e.g., Mediterranean diet index, activity level quartiles).

Model Development

We developed a fully connected feed-forward neural network:

- **Architecture:** Input layer with 25 variables, 3 hidden layers (128, 64, 32 units) with ReLU activation, and dropout (0.3).
- **Output:** Binary classification with sigmoid activation.
- **Training:** Adam optimizer (learning rate 0.001), batch size 64, early stopping based on validation loss.
- **Data split:** 70% training, 15% validation, 15% external testing.

Comparators

Performance was benchmarked against:

- Logistic regression
- Framingham Risk Score [5]
- QRISK3 [14]

Evaluation Metrics

- Discrimination: AUROC, AUPRC
- Calibration: Hosmer–Lemeshow test, calibration plots
- Clinical utility: Decision Curve Analysis (DCA)
- Interpretability: SHAP values to assess feature importance [12].

3. Results

Model Performance

The neural network achieved an AUROC of 0.89 (95% CI, 0.87–0.91), outperforming logistic regression (0.78), Framingham (0.74), and QRISK3 (0.77). AUPRC was also higher for the neural network (0.81 vs. 0.68 for logistic regression). Calibration analysis demonstrated close alignment between predicted and observed risks, with non-significant Hosmer–Lemeshow statistics ($p = 0.42$). For instance, an AUROC improvement from 0.78 (logistic regression) to 0.89 (neural network) implies that approximately 11 more patients per 100 are correctly ranked in risk order, which directly impacts treatment allocation and prevention strategies.

Clinical Utility

Decision Curve Analysis revealed that the neural network provided greater net benefit across all threshold probabilities. At a threshold of 10%, the model avoided 8 unnecessary interventions per 100 patients compared with Framingham and 6 compared with QRISK3.

Interpretability

SHAP analysis identified the most important predictors as:

1. Systolic blood pressure
2. LDL cholesterol
3. Physical activity
4. Family history of CVD
5. Stress score

Interestingly, lifestyle variables such as diet quality and stress, often overlooked in conventional models, ranked higher than traditional risk factors such as age and smoking status [11].

4. Discussion

Our findings demonstrate that neural networks substantially outperform traditional risk prediction models in CVD prediction, consistent with prior evidence in machine learning–driven risk stratification [3,7]. The enhanced predictive accuracy can be attributed to the ability of neural networks to capture nonlinear and interactive effects among risk factors [4].

Importantly, incorporating lifestyle and psychosocial factors significantly improved model performance, reinforcing the multidimensional nature of cardiovascular risk [11]. This aligns with previous calls to integrate non-traditional risk markers into CVD prediction [6,10]. The interpretability provided by SHAP values further supports the clinical acceptability of these models, addressing concerns about the "black box" nature of AI [12].

From a clinical perspective, these models can be integrated into electronic health record systems to provide real-time risk predictions, supporting shared decision-making between clinicians and patients. Moreover, they offer a pathway toward precision cardiology by tailoring preventive interventions based on individualized risk profiles [9,15].

However, challenges remain. Neural networks require large, high-quality datasets for training, raising concerns about generalizability across underrepresented populations [8]. Further, computational demands and infrastructure requirements may limit scalability in resource-constrained settings. External multi-center prospective validations remain essential before broad clinical adoption [13,14].

Future directions include exploring survival-based deep learning models, hybrid ontology-driven neural networks, and federated learning frameworks that allow for privacy-preserving model training across multiple institutions [15]. In addition to generalizability and computational challenges, reliance on EHR-derived data introduces potential biases, such as missingness not at random and underrepresentation of lower socioeconomic groups, which could affect model fairness.

5. Conclusions

This study confirms that neural networks provide superior predictive accuracy, calibration, and clinical utility compared with traditional cardiovascular risk models. By integrating a broader range of variables—including lifestyle and psychosocial factors—these models align with precision medicine principles and offer meaningful improvements for preventive cardiology [1,9,11].

The clinical implications are significant: improved risk prediction can facilitate earlier interventions, optimize resource allocation, and reduce unnecessary treatments [6,13]. With interpretability tools such as SHAP, the clinical adoption barrier is lowered, making neural networks not only powerful but also transparent and trustworthy [12]. The identification of stress and diet as top predictors highlights modifiable targets for preventive interventions, suggesting that integration of mental health support and nutritional counseling into standard CVD prevention could yield meaningful benefits.

While challenges in generalizability, scalability, and external validation remain, neural networks represent a transformative step in cardiovascular risk prediction. Their integration into healthcare systems could shift the paradigm from reactive treatment toward proactive, individualized prevention, ultimately improving population health outcomes [3,10,15].

REFERENCES

[1] Benjamin EJ, et al. Heart Disease and Stroke Statistics—2024 Update. *Circulation*. 2024; 149: e95–e114.

[2] D'Agostino RB, et al. General cardiovascular risk profile for use in primary care. *Circulation*. 2008; 117(6): 743–753.

[3] Esteva A, et al. A guide to deep learning in healthcare. *Nat Med*. 2019; 25(1): 24–29.

[4] Krittanawong C, et al. Artificial intelligence in precision cardiovascular medicine. *Nat Rev Cardiol*. 2021; 18(9): 592–609.

[5] Hippisley-Cox J, et al. Predicting cardiovascular risk in England and Wales: QRISK2. *BMJ*. 2008; 336: 1475–1482.

[6] Yusuf S, et al. Modifiable risk factors, cardiovascular disease, and mortality. *Lancet*. 2004; 364(9438): 937–952.

[7] Shameer K, et al. Machine learning in cardiovascular medicine. *J Am Coll Cardiol*. 2018; 71(23): 2668–2679.

[8] Vickers AJ, et al. Limitations of risk prediction models in diverse populations. *Stat Med*. 2016; 35(9): 1330–1343.

[9] Johnson KW, et al. Artificial intelligence in cardiology. *J Am Coll Cardiol*. 2018; 71(23): 2668–2679.

[10] Ioannidis JPA. Integration of biomarkers into CVD prediction. *JAMA*. 2018; 320(19): 1975–1977.

[11] Lear SA, et al. Lifestyle factors and cardiovascular disease risk. *Eur Heart J*. 2017; 38(22): 1719–1726.

[12] Lundberg SM, et al. Explainable AI for healthcare. *Nat Biomed Eng*. 2020; 4: 566–573.

[13] Collins GS, et al. Transparent reporting of prediction models: TRIPOD statement. *Ann Intern Med*. 2015; 162(1): 55–63.

[14] Hippisley-Cox J, et al. Development and validation of QRISK3 risk prediction algorithm. *BMJ*. 2017; 357: j2099.

[15] Li Y, et al. Federated learning for cardiovascular risk prediction. *J Am Med Inform Assoc*. 2022; 29(3): 340–348.

[16] Bilal, O., Hekmat, A., Shahzad, I., & Raza, A. (2025). Boosting Machine Learning Accuracy for Cardiac Disease Prediction: The Role of Advanced Feature Engineering and Model Optimization. The Review of Socionetwork Strategies. Springer. https://doi.org/10.1007/s12626-025-00190-w.

Copyright © 2025 The Author(s). Published by Scientific & Academic Publishing
This work is licensed under the Creative Commons Attribution International License (CC BY). http://creativecommons.org/licenses/by/4.0/