International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Comparative Analysis of Artificial Intelligence Techniques for Cardiovascular Disease Diagnosis and Risk Prediction

Shaikh Marium Shujauddin[1], Dr.H.S.Fadewar[2]

[1,2]Swami Ramanand Teerth Marathwada University, Nanded, Recognized By the UGC U/s 2(f) and 12(B) NAAC Re-accredited 'B++' grade with CGPA 2.96

ABSTRACT

Cardiovascular disease (CVD) remains the leading cause of global mortality, necessitating accurate and clinically generalizable diagnostic models. Artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), has demonstrated potential in analyzing clinical data for early CVD detection. However, existing studies predominantly focus on single-model performance within isolated datasets, with limited cross-comparative synthesis.

This study provides a comparative evaluation of major AI algorithms for CVD diagnosis and risk prediction. A narrative review of comparative studies published between 2019 and 2026 was conducted using PubMed, Google Scholar, and SciSpace. Performance metrics were synthesized across models including Random Forest (RF), Support Vector Machines (SVM), k-Nearest Neighbors (KNN), Gradient Boosting, ensemble methods, and Deep Neural Networks (DNNs).

Results reveal ensemble and Gradient Boosting approaches consistently demonstrated superior predictive performance (AUC: 0.80–0.92). Random Forest frequently outperformed KNN and logistic regression in structured clinical data. Although deep learning models achieved high accuracy in single-cohort studies, limited external validation constrains clinical applicability.

This study provides evidence that ensemble-based frameworks currently offer the most effective balance between predictive accuracy, robustness, and clinical feasibility. Future research should emphasize multi-center external validation and explainable AI frameworks.

Keywords: Cardiovascular disease; Artificial intelligence; Machine learning; Deep learning; Risk prediction

1. INTRODUCTION

1.1 Background

Cardiovascular disease (CVD) remains the foremost cause of mortality worldwide, claiming 17.9 million lives annually—32% of all global deaths (WHO, 2023). Most deaths occur in low- and middle-income countries with limited access to early diagnosis (Roth et al., 2020). As populations age and risk factors like hypertension and diabetes rise, CVD burden will increase substantially (Virani et al., 2021).

Early diagnosis and reliable risk prediction are critical for reducing mortality. Traditional approaches like the Framingham Risk Score (Wilson et al., 1998) and SCORE (Conroy et al., 2003) have proven valuable but are limited by restricted risk factors and assumptions of linear relationships (D'Agostino et

al., 2008). These models often underperform in diverse populations (DeFilippis et al., 2015).

1.2 AI in Cardiovascular Medicine

Healthcare data growth—electronic health records (EHRs), imaging, genomics, wearables—has positioned AI as transformative in cardiovascular medicine (Krittanawong et al., 2020). Machine learning (ML) and deep learning (DL) identify complex, non-linear relationships in high-dimensional data and integrate diverse modalities (Shameer et al., 2018).

Techniques like Random Forest (RF), Support Vector Machines (SVM), and k-Nearest Neighbors (KNN) have been extensively applied to structured clinical data (Weng et al., 2017). Ensemble methods including Gradient Boosting (GBM, XGBoost) combine weak learners into robust models (Zhang et al., 2021). Deep learning approaches—Deep Neural Networks (DNNs) and Convolutional Neural Networks (CNNs)—show success with imaging and ECG data (Attia et al., 2019).

1.3 Current State of Research

Weng et al. (2017) found RF and Gradient Boosting achieved AUC >0.75. Ambale-Venkatesh et al. (2017) demonstrated improved prediction using ML. Kwon et al. (2020) developed DNN models with AUC 0.86. Hannun et al. (2019) showed DL detected atrial fibrillation exceeding cardiologist accuracy.

However, most studies focus on single-model performance within isolated datasets (Liu et al., 2022), making it difficult to determine which techniques are most appropriate for specific clinical contexts.

1.4 The Gap

Reviews often focus on single algorithmic categories or provide narrative overviews without systematic comparison (Johnson et al., 2021). Existing literature insufficiently addresses generalizability and clinical applicability. Algorithms may excel in single-cohort studies but fail in diverse populations (Kelly et al., 2019). External validation—prerequisite for clinical deployment—remains poorly characterized (Collins et al., 2021).

1.5 Study Objectives

This review provides comparative evaluation of major AI algorithms for CVD diagnosis and risk prediction, examining:

1. Performance metrics across RF, SVM, KNN, Gradient Boosting, ensemble methods, and DNNs
2. Algorithms demonstrating superior performance across contexts
3. Validation practices and generalizability
4. Clinical applicability factors beyond raw performance
5. Evidence-based guidance for model selection

2. METHODOLOGY

2.1 Study Design

This narrative review employed comparative synthesis to examine AI techniques for CVD diagnosis and risk prediction, synthesizing insights from representative studies (Green et al., 2006).

2.2 Research Questions

1. What are reported performance metrics of major AI algorithms for CVD prediction?
2. Which algorithms demonstrate superior performance across datasets?
3. What validation practices are employed?
4. What factors influence clinical applicability?

2.3 Search Strategy

Databases searched: PubMed/MEDLINE, Google Scholar, SciSpace (January 2019–March 2026).

Search terms: "cardiovascular disease" OR "CVD" AND "machine learning" OR "deep learning" AND "comparative study" OR "comparison."

2.4 Inclusion Criteria

- Human subjects for CVD diagnosis/risk prediction
- At least two AI algorithms compared
- Reporting of accuracy, sensitivity, specificity, and/or AUC
- Peer-reviewed original research, English, 2019–2026

2.5 Exclusion Criteria

- Single-algorithm studies
- Conference abstracts, editorials, preprints
- Non-CVD outcomes
- Insufficient metrics
- Reviews without original data

2.6 Study Selection

Six representative comparative studies were selected for in-depth review, representing diversity in algorithms, data sources, outcomes, and geographical settings (Ferrari, 2015).

2.7 Data Extraction

Extracted: author, year, population, sample size, data source, algorithms, performance metrics, validation approach, key findings.

2.8 Data Synthesis

Narrative synthesis focused on algorithmic performance patterns, validation practices, and clinical applicability (Popay et al., 2006). Meta-analysis was not performed due to heterogeneity.

2.9 Limitations

This review acknowledges selection bias, small sample (six studies), language restriction, and no formal quality assessment.

3. RESULTS

3.1 Overview

Six studies published 2019–2026 were analyzed, encompassing diverse algorithms and datasets.

Table 1

Comparative Performance of AI Techniques for Cardiovascular Disease Diagnosis

Algorithm Class	Reported Performance	Key Findings
Random Forest	Superior to KNN, logistic regression	Robust with high-dimensional data; strong in structured clinical environments [1]
Ensemble/Gradient Boosting	AUC: 0.80–0.92	Most consistent across datasets; excellent stability [2]
Shallow ANN	Accuracy: 88.5%, Recall: 90%	Good balance of sensitivity/specificity; computationally efficient [3]

Algorithm Class	Reported Performance	Key Findings
Deep Learning (DNN/CNN)	Up to 99.99% (Cleveland)	High single-cohort accuracy but limited external validation [4]
SVM/KNN	Variable, below ensemble methods	Dataset-dependent; baseline comparators [1,5]
Graph Neural Networks	Promising, preliminary	Effective for pathway integration; needs validation [2]

3.2 Performance Patterns

Random Forest: Consistently outperformed KNN and logistic regression due to ensemble nature reducing overfitting through bootstrap aggregation [1].

Ensemble/Gradient Boosting: Most consistent performers (AUC 0.80–0.92). Sequential learning corrects predecessor errors, balancing bias and variance effectively [2].

Shallow ANN: Accuracy 88.5% with 90% recall. Middle ground between classical ML and deep learning; superior interpretability to deep networks [3].

Deep Learning: DenseNet achieved 99.99% on Cleveland dataset (n=303). However, single-cohort design raises overfitting concerns; no external validation reported [4].

SVM/KNN: Consistent underperformance relative to ensemble methods. SVM sensitive to kernel selection; KNN degrades in high-dimensional spaces [1,5].

Graph Neural Networks: Promising for integration but limited by small cohorts and interpretability challenges [2].

3.3 Validation Practices

All studies reported internal validation (cross-validation, train-test split). No study reported external validation on independent datasets. This represents the greatest barrier to clinical translation.

3.4 Algorithmic Trade-offs

- Interpretability vs. Accuracy: RF/Gradient Boosting offer feature importance; deep learning operates as "black boxes"
- Data Efficiency: Ensemble methods perform well with thousands of patients; deep learning requires tens of thousands
- Computational Requirements: Ensemble methods train on standard hardware; deep learning needs GPUs
- Reproducibility: Extreme accuracy studies often lack methodological detail

4. DISCUSSION

4.1 Key Findings

Ensemble and Gradient Boosting methods demonstrated most consistent performance (AUC 0.80–0.92). Their sequential learning mechanism suits CVD's incremental risk accumulation (Dietterich, 2000).

Random Forest consistently outperformed SVM/KNN in structured data, aligning with Weng et al. (2017). Feature importance aids clinical interpretation.

Deep learning achieved highest accuracy in single cohorts but lacks external validation. The Cleveland dataset (n=303) achieving 99.99% raises overfitting concerns despite regularization (Kelly et al., 2019). Validation practices are insufficient. No study reported external validation, though this is essential for clinical deployment (Collins et al., 2021).

4.2 Clinical Implications

Accuracy vs. Generalizability: Ensemble methods offer stability across contexts; deep learning requires external validation before deployment.

Interpretability: RF and Gradient Boosting provide SHAP values and feature importance—critical for clinician trust (Cutillo et al., 2020).

Data Efficiency: Ensemble methods accessible to institutions with limited data; deep learning requires large consortia.

Computational Requirements: Ensemble methods deployable in resource-constrained settings; deep learning needs specialized hardware.

4.3 The Validation Crisis

Only 6% of medical AI studies perform external validation (Kim et al., 2019). Models fail due to:

- Population shift
- Measurement shift
- Practice shift
- Temporal shift

Without multi-center validation, models remain experimental (Benjamens et al., 2020).

4.4 Future Research Priorities

1. Prioritize external validation through multi-center collaborations
2. Establish standardized benchmarking protocols
3. Integrate explainable AI from development outset
4. Address health equity through diverse representation
5. Develop implementation science approaches

4.5 Limitations

- Small sample (six studies)
- Narrative design (no meta-analysis)
- Selection and publication bias possible
- Language restriction to English
- Rapidly evolving field

5. CONCLUSION

5.1 Summary

This review provides comparative evaluation of AI techniques for CVD prediction:

1. Ensemble/Gradient Boosting demonstrate most consistent performance (AUC 0.80–0.92)—optimal balance of accuracy and generalizability
2. Random Forest consistently outperforms classical ML algorithms
3. Deep learning achieves high accuracy but lacks external validation
4. Critical validation gap exists across literature
5. Algorithm selection involves accuracy-interpretability trade-offs

5.2 Recommendations

For Researchers:

- Prioritize external validation
- Adopt standardized reporting
- Integrate explainable AI
- Ensure diverse representation
- Share code for replication

For Clinicians:

- View single-cohort studies cautiously
- Prioritize externally validated algorithms
- Consider interpretability alongside accuracy

For Journals:

- Require external validation evidence
- Enforce reporting standards
- Encourage negative results publication

For Regulators:

- Maintain rigorous validation standards
- Require diverse population evidence

5.3 Final Remarks

Ensemble methods currently offer the most reliable pathway to clinical deployment. Deep learning's role depends on demonstrating generalizability. Collaborative effort across researchers, clinicians, and regulators is essential to fulfill AI's potential in reducing global CVD burden.

REFERENCES

1. Ambale-Venkatesh, B., Yang, X., Wu, C. O., Liu, K., Hundley, W. G., McClelland, R., Gomes, A. S., Folsom, A. R., Shea, S., Guallar, E., Bluemke, D. A., & Lima, J. A. C. (2017). Cardiovascular event prediction by machine learning: The Multi-Ethnic Study of Atherosclerosis. Circulation Research, 121(9), 1092–1101.
2. Attia, Z. I., Noseworthy, P. A., Lopez-Jimenez, F., Asirvatham, S. J., Deshmukh, A. J., Gersh, B. J., Carter, R. E., Yao, X., Rabinstein, A. A., Erickson, B. J., Kapa, S., & Friedman, P. A. (2019). An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm. The Lancet, 394(10201), 861–867.
3. Benjamens, S., Dhunnoo, P., & Meskó, B. (2020). The state of artificial intelligence-based FDA-approved medical devices and algorithms. NPJ Digital Medicine, 3(1), Article 118.
4. Beunza, J. J., Puertas, E., García-Ovejero, E., Villalba, G., Condés, E., Koleva, G., Hurtado, C., & Landecho, M. F. (2019). Comparison of machine learning algorithms for clinical event prediction. Journal of Biomedical Informatics, 97, Article 103257. [1]
5. Choudhury, B. (2023). Automated cardiovascular disease prediction models: A comparative analysis. EAI Endorsed Transactions on Pervasive Health and Technology, 9, e6. [2]
6. Collins, G. S., Dhiman, P., Andaur Navarro, C. L., Ma, J., Hooft, L., Reitsma, J. B., Logullo, P., Beam, A. L., Peng, L., Van Calster, B., van Smeden, M., Riley, R. D., & Moons, K. G. M. (2021). Protocol for development of a reporting guideline (TRIPOD-AI). BMJ Open, 11(7), Article e048008.
7. Conroy, R. M., Pyörälä, K., Fitzgerald, A. P., Sans, S., Menotti, A., De Backer, G., De Bacquer, D.,

Ducimetière, P., Jousilahti, P., Keil, U., Njølstad, I., Oganov, R. G., Thomsen, T., Tunstall-Pedoe, H., Tverdal, A., Wedel, H., Whincup, P., Wilhelmsen, L., & Graham, I. M. (2003). Estimation of ten-year risk of fatal cardiovascular disease in Europe: The SCORE project. European Heart Journal, 24(11), 987–1003.

8. Cutillo, C. M., Sharma, K. R., Foschini, L., Kundu, S., Mackintosh, M., & Mandl, K. D. (2020). Machine intelligence in healthcare—Perspectives on trustworthiness, explainability, usability, and transparency. NPJ Digital Medicine, 3(1), Article 47.
9. D'Agostino, R. B., Sr., Vasan, R. S., Pencina, M. J., Wolf, P. A., Cobain, M., Massaro, J. M., & Kannel, W. B. (2008). General cardiovascular risk profile for use in primary care. Circulation, 117(6), 743–753.
10. DeFilippis, A. P., Young, R., Carrubba, C. J., McEvoy, J. W., Budoff, M. J., Blumenthal, R. S., Kronmal, R. A., McClelland, R. L., Nasir, K., & Blaha, M. J. (2015). An analysis of calibration and discrimination among multiple cardiovascular risk scores. Annals of Internal Medicine, 162(4), 266–275.
11. Dietterich, T. G. (2000). Ensemble methods in machine learning. Lecture Notes in Computer Science, 1857, 1–15.
12. Ferrari, R. (2015). Writing narrative style literature reviews. Medical Writing, 24(4), 230–235.
13. Green, B. N., Johnson, C. D., & Adams, A. (2006). Writing narrative literature reviews for peer-reviewed journals. Journal of Chiropractic Medicine, 5(3), 101–117.
14. Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C., Turakhia, M. P., & Ng, A. Y. (2019). Cardiologist-level arrhythmia detection using a deep neural network. Nature Medicine, 25(1), 65–69.
15. Johnson, K. W., Torres Soto, J., Glicksberg, B. S., Shameer, K., Miotto, R., Ali, M., Ashley, E., & Dudley, J. T. (2021). Artificial intelligence in cardiology. Journal of the American College of Cardiology, 77(14), 1758–1771.
16. Kandula, S. T. R. (2025, July). Comparison and performance assessment of intelligent ML models for forecasting cardiovascular disease risks [Conference paper]. 2025 International Conference on Sustainable Energy Networks and Smart Technologies, 1–6. [3]
17. Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. BMC Medicine, 17(1), Article 195.
18. Kim, D. W., Jang, H. Y., Kim, K. W., Shin, Y., & Park, S. H. (2019). Design characteristics of studies reporting the performance of AI algorithms for medical image analysis. Korean Journal of Radiology, 20(3), 405–410.
19. Krittanawong, C., Johnson, K. W., Rosenson, R. S., Wang, Z., Aydar, M., Baber, U., Min, J. K., Tang, W. H. W., Halperin, J. L., & Narayan, S. M. (2020). Deep learning for cardiovascular medicine: A practical primer. European Heart Journal, 41(21), 1988–1999.
20. Kwon, J. M., Lee, Y., Lee, Y., Lee, S., & Park, J. (2020). Deep-learning algorithm for electrocardiography-based heart failure identification. Korean Circulation Journal, 50(7), 582–593.
21. Liu, X., Faes, L., Kale, A. U., Wagner, S. K., Fu, D. J., Bruynseels, A., Mahendiran, T., Moraes, G., Shamdas, M., Kern, C., Ledsam, J. R., Schmid, M. K., Balaskas, K., Topol, E. J., Bachmann, L. M., Keane, P. A., & Denniston, A. K. (2022). Deep learning performance vs. health-care professionals in detecting diseases from medical imaging. The Lancet Digital Health, 4(1), e51–e66.
22. Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., &

Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. ESRC Methods Programme, 1(1), b92.

23. Roth, G. A., Mensah, G. A., Johnson, C. O., Addolorato, G., Ammirati, E., Baddour, L. M., Barengo, N. C., Beaton, A. Z., Benjamin, E. J., Benziger, C. P., Bonny, A., Brauer, M., Brodmann, M., Cahill, T. J., Carapetis, J., Catapano, A. L., Chugh, S. S., Cooper, L. T., Coresh, J., ... Fuster, V. (2020). Global burden of cardiovascular diseases and risk factors, 1990–2019. Journal of the American College of Cardiology, 76(25), 2982–3021.
24. Shameer, K., Johnson, K. W., Glicksberg, B. S., Dudley, J. T., & Sengupta, P. P. (2018). Machine learning in cardiovascular medicine: Are we there yet? Heart, 104(14), 1156–1164.
25. Udoy, M. (2022). Performance evaluation and comparative analysis of different machine learning algorithms in predicting cardiovascular disease [Unpublished manuscript]. [4]
26. Virani, S. S., Alonso, A., Aparicio, H. J., Benjamin, E. J., Bittencourt, M. S., Callaway, C. W., Carson, A. P., Chamberlain, A. M., Cheng, S., Delling, F. N., Elkind, M. S. V., Evenson, K. R., Ferguson, J. F., Gupta, D. K., Khan, S. S., Kissela, B. M., Knutson, K. L., Lee, C. D., Lewis, T. T., ... Tsao, C. W. (2021). Heart disease and stroke statistics—2021 update. Circulation, 143(8), e254–e743.
27. Weng, S. F., Reps, J., Kai, J., Garibaldi, J. M., & Qureshi, N. (2017). Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS ONE, 12(4), Article e0174944.
28. Wilson, P. W., D'Agostino, R. B., Levy, D., Belanger, A. M., Silbershatz, H., & Kannel, W. B. (1998). Prediction of coronary heart disease using risk factor categories. Circulation, 97(18), 1837–1847.
29. World Health Organization. (2023). Cardiovascular diseases (CVDs). https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)
30. Yan, C., Xing, Y., Liu, S., Gao, E., & Wang, J. (2024). Machine learning models for cardiovascular disease prediction: A comparative study [Preprint]. *medRxiv. [5]
31. Yazdi, N. A., Noori, M., Ayyoubzadeh, S. M., Naghdi, S., & Saeedi, S. (2025). Predicting triage levels in patients presenting with cardiac-related symptoms: A comparison of supervised machine learning methods. BMC Emergency Medicine, 25(1), Article 45. [6]
32. Zhang, Z., Ho, K. M., & Hong, Y. (2021). Machine learning for the prediction of cardiovascular disease: A systematic review and meta-analysis. Journal of the American Medical Informatics Association, 28(5), 1018–1028.