International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 12 Issue: 10 | Oct 2025 www.irjet.net p-ISSN: 2395-0072

Multimodal Deep Learning Algorithms for Predictive Modeling of Cardiovascular Disease Onset and Progression: A Tetramodal Architecture Integrating HER, Imaging, Genomics, and Wearables Data

Dr. Emmanuel Ameh[1], Dr. Alexander Paselk[2]

Dept. Machine Learning, Capitol Technology University, Maryland, USA
Dept. Occupational Health and Safety, Capitol Technology University, Maryland, USA

Abstract - *Background: Cardiovascular diseases (CVDs) remain the leading global cause of mortality, accounting for over 18 million deaths annually. Traditional risk scores, such as the Framingham Risk Score, fail to fully capture the complex interactions between genetic, physiological, and lifestyle factors. Deep learning methods, particularly multimodal architectures, offer opportunities to integrate heterogeneous data sources for improved prediction of both CVD onset and progression. Methods: This retrospective, quantitative, cohort study developed and evaluated a tetramodal deep learning framework integrating electronic health records (EHR), imaging, genomic, and wearables data. Individual modality architectures included convolutional neural networks (CNNs) for imaging, transformer encoders for wearables, and fully connected networks for EHR and genomics. Early fusion integration was employed to construct the tetramodal architecture. Data preprocessing, feature extraction, and multimodal fusion strategies were optimized via ablation studies. Model performance was evaluated for binary classification (CVD_Presence) and regression (Time_to_Event) tasks using AU-ROC, F1-score, mean absolute error (MAE), and the Friedman chi-squared test with post-hoc analysis. Results: The tetramodal model significantly outperformed all single-, bi-, and tri-modal models for both CVD onset and progression prediction. Inclusion of wearables data improved predictive accuracy for progression tasks with statistical significance ($p = 0.04$). The architecture demonstrated superior sensitivity, specificity, and clinical applicability compared to unimodal baselines. Conclusions: This study presents a scalable, interpretable tetramodal deep learning framework capable of integrating heterogeneous clinical and real-time physiological data for enhanced CVD risk stratification. Future work will focus on real-world deployment, optimization for low-resource environments, and external validation on diverse patient populations.*

Key Words: Deep Learning Models, Cardiovascular Disease Prediction, Machine Learning Models, Multimodal Data Modeling, Tetramodal Data, Wearable Data

1.INTRODUCTION

Cardiovascular diseases (CVDs) constitute a persistent public health burden, causing approximately one-third of global deaths annually (WHO, 2024). Despite advancements in clinical care, risk prediction remains dominated by traditional models such as the Framingham Risk Score, which rely on a limited set of clinical variables (cholesterol, age, smoking status) and fail to leverage the breadth of modern patient data sources. These limitations contribute to misclassification of high-risk patients and insufficient monitoring of disease progression, particularly in those with subclinical presentations.

Recent advances in deep learning (DL) enable the processing of high-dimensional, multimodal datasets, ranging from imaging and genomic sequences to continuous physiological monitoring from wearable devices. Multimodal deep learning, specifically tetramodal integration in this study, addresses the complexity of CVD pathophysiology by fusing heterogeneous data modalities for richer, more personalized risk assessment. The objective of this study was to design, implement, and evaluate a tetramodal deep learning architecture for predicting both CVD onset (binary classification: CVD_Presence) and CVD progression (regression: Time_to_Event).

Despite its promise, the application of multimodal deep learning to CVD prediction remains limited. Persistent challenges include the interpretability of complex models in clinical settings, the computational cost of training and deployment, integration of longitudinal patient data, and variability in data formats and quality (Abbas & Daena, 2025; Nadella et al., 2023). Furthermore, the absence of standardized algorithms for multimodal integration and reliance on small-scale datasets reduce clinical translatability (Abbas & Daena, 2025; Esteva et al., 2021). Research on leveraging multimodal deep learning for predicting CVD progression, especially in patients with multiple comorbidities, remains scarce (Kasula, 2023; Soenksen et al., 2022; Vaid et al., 2023).

Addressing these gaps requires robust, generalizable multimodal deep learning algorithms capable of integrating heterogeneous data streams to predict both onset and progression of CVD. Accurate, timely predictions could facilitate earlier interventions, reduce complications, and improve patient outcomes (V. V. Paul & Masood, 2024; Terranova & Venkatakrishnan, 2024; Tian et al., 2024). The present study develops and evaluates such a framework,

aiming to transition from unimodal research models to clinically viable, real-world multimodal predictive systems capable of reducing the global burden of cardiovascular disease.

2. BACKGROUND OF THE STUDY

Cardiovascular diseases (CVDs) encompass a broad spectrum of conditions, including hypertension, stroke, heart failure, and coronary artery disease, driven by a complex interplay of environmental, genetic, and lifestyle factors (Huang et al., 2023; Kasula, 2023; Xia et al., 2024; Y. Zhou et al., 2024). Despite advances in diagnostics and therapeutics, predicting CVD onset and progression remains challenging due to the multifactorial nature of these conditions. Early detection is crucial for prevention and management, and personalized medicine plays a pivotal role in improving outcomes (Tomov & Tomov, 2021).

2.1 Emergence of Multimodal Deep Learning in Healthcare

Deep learning has transformed fields such as natural language processing, image recognition, and healthcare (Armoundas et al., 2024; Esteva et al., 2021; Joshi et al., 2024), enabling automatic learning of hierarchical representations from raw data without the need for manual feature engineering (LeCun et al., 2015; C. Zhou et al., 2024). However, most deep learning applications process single data modalities, whereas modern clinical practice routinely involves multimodal data, ranging from structured clinical records to sensor streams (Naser et al., 2024; Skorupko et al., 2024; Subha & Kasturi, 2024).

2.2 Challenges and Opportunities

Key barriers to adoption include data heterogeneity, requiring harmonization across formats and sampling rates (F. Ali et al., 2020; Mohite et al., 2024; Patel et al., 2024), high computational demands (Adelson et al., 2023; Nomura et al., 2025; Rijken et al., 2025), and limited interpretability (C. C. Yang, 2022; Tuli et al., 2020; Vaid et al., 2023). Explainability tools such as LIME and SHAP are increasingly used to elucidate feature contributions (Karna et al., 2024; Ullah & Garcia-Zapirain, 2024; Xia et al., 2024). Meanwhile, wearable integration offers novel opportunities for continuous, personalized monitoring (K. Khan et al., 2025; Botros et al., 2025; Sikdar et al., 2025). This study develops and evaluates multimodal deep learning algorithms integrating imaging, EHR, genomic, and wearable data, aiming to enhance predictive accuracy for both CVD onset and progression while addressing data integration, computational, and interpretability challenges.

3. PURPOSE OF THE STUDY

This quantitative, non-experimental cohort study aimed to design, develop, and validate efficient multimodal deep learning algorithms for accurate predictive modeling of cardiovascular disease (CVD) onset and progression. Given the global burden of CVD, with high mortality and morbidity rates, enhancing the precision and timeliness of predictive models is critical for optimizing preventive strategies, enabling personalized interventions, and improving patient outcomes (Ravi & Madhavan, 2024; WHO, 2024; C. Zhou et al., 2024). Current CVD prediction tools, such as traditional risk scores and conventional statistical models, are limited in their capacity to capture the multifactorial nature of cardiovascular pathology. These models rely on narrow sets of clinical variables and fail to account for complex interactions among genetic, physiological, and lifestyle factors. Advances in deep learning, particularly architectures capable of integrating heterogeneous data sources, offer a pathway toward more robust, individualized predictions (Gulshan et al., 2016; Joshi et al., 2024; Vaid et al., 2023). Finally, this work sought to empirically evaluate whether multimodal deep learning approaches significantly outperform unimodal counterparts in predictive accuracy, thereby advancing the evidence base for the use of multimodal AI systems in cardiovascular precision medicine.

3.1 Significance of the Study

Cardiovascular diseases (CVDs), including hypertension, stroke, heart failure, and coronary artery disease, arise from a complex interplay of environmental, genetic, and lifestyle factors (Huang et al., 2023; Joshi et al., 2024; Martin et al., 2025). Despite advances in diagnostics and therapeutics, accurate prediction of CVD onset and progression remains a major challenge due to the multifactorial nature of these conditions. Early detection is essential for prevention and effective management, while personalized medicine is increasingly recognized as critical in reducing CVD-related mortality and morbidity. This study is significant because it advances the application of deep learning in healthcare by focusing on multimodal data integration for cardiovascular disease prediction. Unlike conventional risk models, which are constrained by limited clinical variables, multimodal deep learning leverages diverse data sources, including imaging, genomic, wearable sensor, and structured clinical data to construct richer, more comprehensive models (Gulshan et al., 2016; Raj & Bayappu, 2024; Vaid et al., 2023). By uniting these modalities, the approach can outperform traditional tools in both predictive accuracy and clinical applicability.

3.2 Hypotheses/Research Questions

Research Question 1: What is the comparative predictive performance of multimodal deep learning models versus single-modality models in predicting CVD presence, and how do they differ in terms of sensitivity, accuracy, and specificity? Research Question 2: How does the predictive performance of multimodal deep learning models compare to traditional cardiovascular risk models in predicting cardiovascular disease progression? Research Question 3:

How does the integration of wearables data with genomic, imaging, and clinical data in multimodal deep-learning models improve the real-time prediction of adverse cardiovascular events compared to models without wearables data?

3.3 Research Hypotheses

Research Hypothesis 1: Multimodal deep learning models will outperform single-modality models in predicting CVD presence. $H1_0$: Multimodal deep learning models do not significantly improve the prediction of CVD presence compared to single-modality models. H1a: Multimodal deep learning models significantly improve the prediction of CVD presence compared to single-modality models. Research Hypothesis 2: Multimodal deep learning models can more accurately predict the time to a cardiovascular event compared to traditional risk score models. $H2_0$: Multimodal deep learning models do not outperform traditional risk models in predicting the time to a cardiovascular event. H2a: Multimodal deep learning models significantly outperform traditional risk models in predicting the time to a cardiovascular event. Research Hypothesis 3: The integration of wearable data in multimodal deep learning models will significantly improve the real-time prediction of adverse cardiovascular events. $H3_0$: The integration of wearables data does not significantly improve real-time prediction of adverse cardiovascular events in multimodal deep learning models.

H3a: The integration of wearables data significantly improves real-time prediction of adverse cardiovascular events in multimodal deep learning models.

4. THEORETICAL FRAMEWORK

This research integrates technological, computational, and biomedical foundations to inform the predictive modeling of cardiovascular disease (CVD) onset and progression. The framework draws on three core pillars: (1) deep learning architectures for heterogeneous data processing (Azevedo et al., 2024; Dosovitskiy et al., 2021), (2) multimodal data integration strategies (Cao et al., 2019; Kanchanamala et al., 2023), and (3) cardiovascular pathophysiology (Jaltotage et al., 2024). CVDs are inherently multifactorial, involving an interplay of environmental, genetic, and behavioral factors that manifest through diverse mechanisms, including atherosclerosis, endothelial dysfunction, and arrhythmogenesis (Botros et al., 2025; Sikdar et al., 2025). This complexity necessitates predictive models capable of incorporating variables spanning wearable device metrics (e.g., physical activity, heart rate variability), genomic risk profiles (e.g., SNPs, polygenic scores), clinical measurements (e.g., cholesterol, blood pressure), imaging findings (e.g., coronary calcium scores, echocardiograms), and demographic factors (Karna et al., 2024; Xia et al., 2024). The theoretical premise is that integrating these modalities offers a more comprehensive assessment of CVD risk and progression than unimodal approaches (Gulshan et al., 2016; Vernooij et al., 2023).

5. ASSUMPTION

The validity, reliability, and clinical applicability of this research are underpinned by several key assumptions related to data quality, integration, model design, interpretability, and generalizability. These assumptions provide the foundation for both the methodological rigor and the robustness of the study's findings.

5.1 Data Harmonization and Standardization

It is assumed that all data, whether sourced from wearable devices, genomic repositories, imaging systems, or clinical records, are accurate, complete, and representative of the patient's actual health status (Dami & Yahaghizadeh, 2021). The predictive performance of deep learning models depends heavily on the integrity of input data (Ajagbe & Adigun, 2024; Das et al., 2022). Inaccuracies such as sensor noise in wearable data or missing values in EHR records can distort learned patterns and degrade predictive accuracy (Udegbe et al., 2024; Yin et al., 2024). Conversely, high-quality, validated datasets improve model robustness and reduce the likelihood of biased predictions (Al-Makhadmeh & Tolba, 2019; Chinni & Manlhiot, 2024).

5.2 Data Accuracy and Completeness

The integration of multimodal data assumes that all modalities, wearable, genomic, imaging, and clinical, are harmonized in structure and standardized in format to enable seamless fusion (Almazroi, 2022; Chicco & Jurman, 2020). Standardization processes are essential for aligning variables across disparate datasets (Sui & Calhoun, 2019; K. Zhang et al., 2023; Y. F. Zhang et al., 2024). For example, spatial alignment of imaging data with clinical measurements, and temporal synchronization of wearable time-series with lab results, are presumed to be achievable and accurate (Amirahmadi et al., 2023; Chafai et al., 2024). Without such harmonization, the model's capacity to uncover cross-modal correlations would be diminished.

5.3 Model Complexity and Suitability

It is assumed that the selected architectures, such as Convolutional Neural Networks (CNNs) for imaging and Long Short-Term Memory (LSTM) networks for longitudinal clinical or wearable data, possess sufficient complexity to capture the nuanced interactions within each modality (Ayano et al., 2023; Bleijendaal et al., 2023). CNNs are presumed to be effective at extracting spatial features from cardiac imaging (LeCun et al., 2015; Li et al., 2021), while LSTMs are assumed to model temporal dependencies in time-series data accurately (Duan et al., 2024). The alignment between data complexity and model capacity is critical to

achieving high predictive accuracy (Bali & Mansotra, 2024; Biswas et al., 2021).

6. HISTORICAL OVERVIEW

6.1 Early Statistical and Unimodal Approaches in Cardiovascular Prediction

Historically, cardiovascular disease (CVD) risk prediction in healthcare relied heavily on unimodal statistical models that drew upon structured clinical datasets. Early methods such as logistic regression and support vector machines (SVMs) were widely employed for their interpretability and ability to handle tabular data (Sadeghi et al., 2024; Y. F. Zhang et al., 2024). These models often required manually engineered features derived from expert clinical knowledge, which limited their scalability and adaptability to new data sources. The Framingham Risk Score (FRS) (D'Agostino et al., 2008) was among the earliest standardized tools, estimating a patient's 10-year risk based on age, cholesterol, blood pressure, diabetes status, and smoking history. Despite its widespread adoption, FRS has been shown to over- or underestimate risk in underrepresented populations (H. Yang et al., 2024) and cannot incorporate richer modalities such as genomic profiles or imaging data. Similarly, the European SCORE (Conroy et al., 2003) and QRISK3 (Hippisley-Cox, 2017) models expanded the variable set to include comorbidities and demographic risk factors but still remained unimodal, failing to integrate imaging, genetic, or continuous wearable data.

6.2 Rise of Classical Machine Learning Models

The 2000s and early 2010s saw the introduction of tree-based ensemble methods such as Random Forests and Gradient Boosting Machines, which could capture non-linear relationships and interactions without explicit feature engineering. Random Forests, in particular, proved effective in high-dimensional EHR datasets, as demonstrated by Khozeimeh et al. (2022) for heart failure prediction. However, these models were not inherently designed for heterogeneous, multimodal data such as medical images, genomic sequences, and time-series wearable streams. Integration required manual feature extraction and concatenation, which constrained scalability. Support Vector Machines (SVMs) also played a major role, especially in small-to-medium structured datasets where high-dimensional separation was advantageous. While offering good performance, SVMs were computationally expensive for very large datasets and unsuitable for raw signal or pixel data without substantial preprocessing.

6.3 The Deep Learning Shift: Automated Feature Learning

The 2010s marked a decisive transition toward deep learning architectures, which automatically learn hierarchical feature representations from raw inputs. This shift eliminated much of the manual feature engineering bottleneck and opened the door to handling unstructured data like images, waveforms, and free text. Convolutional Neural Networks (CNNs) revolutionized medical imaging, enabling accurate segmentation and classification of cardiovascular structures from MRI, CT, and echocardiography (LeCun et al., 2015; Gulshan et al., 2016; Akinola et al., 2024). W. Zhang et al. (2019) demonstrated CNN-based myocardial infarction risk prediction from CCTA, outperforming radiologist assessments. Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRUs), excelled in temporal modeling of ECGs and longitudinal clinical data (Dhaka & Nagpal, 2023; Yao et al., 2020). These models captured disease progression dynamics absent from static risk scores. Although deep learning significantly improved single-modality performance, most early applications still processed one data type at a time. Therefore, limiting their ability to exploit cross-modal interactions inherent in complex diseases like CVD.

6.4 Research Gaps

The predictive modeling of cardiovascular disease (CVD) using multimodal data leverages diverse and heterogeneous sources, including lifestyle metrics, medical imaging, genomic data, and structured clinical records (Ajagbe & Adigun, 2024; Das et al., 2022). While substantial advancements have been made in integrating these sources into deep learning (DL) frameworks, several persistent gaps hinder clinical adoption (Almazroi, 2022; Dami & Yahaghizadeh, 2021). This work aims to address these limitations using state-of-the-art multimodal deep learning architectures. Four primary research gaps were identified through the literature review as follows: i) Lack of Data Integration and Harmonization, ii) Lack of Standardized Data, iii) Computational Complexity, and iv) Model Interpretability and Clinical Trust. Addressing these gaps demands cross-disciplinary collaboration between data scientists, clinicians, biomedical engineers, and policymakers. Key strategies include the adoption of universal healthcare data standards to enable interoperability. It also demands the development of scalable multimodal integration pipelines optimized for clinical latency requirements. Deployment of resource-efficient architectures for equitable access across healthcare settings is also required. Finally, the integration of modality-aware interpretability tools to satisfy regulatory and clinical transparency demands. This research contributes by proposing tetramodal deep learning architectures capable of handling heterogeneous cardiovascular datasets, optimizing computation, and providing interpretable predictions, thereby narrowing the translational gap between research and practice.

7. METHODOLOGY

7.1 Study Design

Given that this study's aim was to build generalizable models for cardiovascular disease (CVD) onset and progression prediction using multimodal deep learning, a quantitative methodology was adopted. The analysis relied heavily on computational modeling and statistical evaluation, consistent with prior CVD prediction studies, all of which employ quantitative designs. A retrospective, quantitative, non-experimental cohort design was employed. The study integrated four primary data modalities: EHR, imaging, genomics, and wearables, collected or simulated to represent a cohort of 50,000 patients. Data preprocessing, model training, and statistical analyses were conducted in Python (TensorFlow, PyTorch) with GPU acceleration.

7.2 Data Modalities

Electronic Health Records (EHR), which are structured clinical features, included demographics (age, gender, ethnicity), comorbidities, medication history, laboratory results (lipid profile, HbA1c, CRP), blood pressure measurements, etc. Imaging data, such as Cardiac CT-derived coronary artery calcium scores and echocardiographic measurements, were extracted. CNNs were applied for spatial feature extraction. Genomic data such as polygenic risk scores (PRS) and selected single-nucleotide polymorphisms (SNPs) relevant to CVD risk were encoded as structured input vectors. Wearables data, which are continuous time-series data from simulated smartwatch devices, included heart rate variability, resting heart rate, physical activity levels, and sleep metrics. Transformer encoders were applied for temporal pattern recognition.

7.3 Synthetic Data Generation and Preprocessing

All datasets were synthetically generated using Python-based Generative Pre-trained Transformer (GPT) modeling to avoid privacy issues and ensure complete multimodal representation. Data was securely stored on the researcher's primary machine with a planned retention of at least three years. Preprocessing involved modality-specific transformation pipelines to standardize formats, normalize scales, and ensure compatibility (Ogunpola et al., 2024; Rahman et al., 2024; Sahoo et al., 2021). Since the data was synthetically generated, no missing values were present, simplifying integration. However, preserving inter-modality relationships remained critical; thus, a multimodal data fusion strategy was applied (Huang et al., 2024; Thakur et al., 2023; Vaid et al., 2023). The preprocessing workflow leveraged Pandas for data wrangling, SciPy for mathematical transformations and imputation when required, and TensorFlow for deep learning model development.

7.4 Data Integration (Fusion)

The study explored multiple fusion strategies to include early fusion, late fusion, and intermediate fusion, as defined in multimodal deep learning literature (Neri et al., 2023; Rehman et al., 2024; Rijken et al., 2025). Ultimately, early fusion was chosen for the final model, as it allowed simultaneous cross-modal feature learning at the input stage, leveraging the complementary nature of heterogeneous modalities.

7.5 Model Development

The deep learning architectures were task-modality matched. Convolutional Neural Networks (CNNs) were selected for imaging data analysis. Multimodal Transformers were chosen for wearable sensor time-series data (Terranova et al., 2024). Multi-Layer Perceptrons (MLPs) with fully connected layers were selected for structured genomic and EHR datasets. The development process followed a structured pipeline, beginning with independent unimodal models trained for each data type, followed by bimodal, trimodal, and tetramodal architectures that were constructed and evaluated. Hyperparameter tuning and optimization was also conducted for each model variant.

7.6 Model Evaluation and Validation

Evaluation incorporated both classification and regression metrics. Classification Metrics included Accuracy, Precision, Recall, F1-score, and Area Under the Receiver Operating Characteristic Curve (ROC-AUC). The Regression Metrics included Concordance Index, Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). Validation followed Stratified K-Fold Cross-Validation to assess generalizability. Statistical significance testing was also employed using the Friedman's Chi-Squared Test, Wilcoxon Signed-Rank Test or Paired t-Test, depending on distributional assumptions (Sirapangi & Gopikrishnan, 2024; Verma et al., 2021).

7.7 Data Preprocessing

Missing data was handled using mode imputation for categorical features and mean imputation for continuous features. The Z-score normalization was applied for EHR and genomics data, whilst min-max scaling was applied for imaging and wearables data. To achieve feature encoding, one-hot encoding was employed for categorical features, and sequence padding for time-series data.

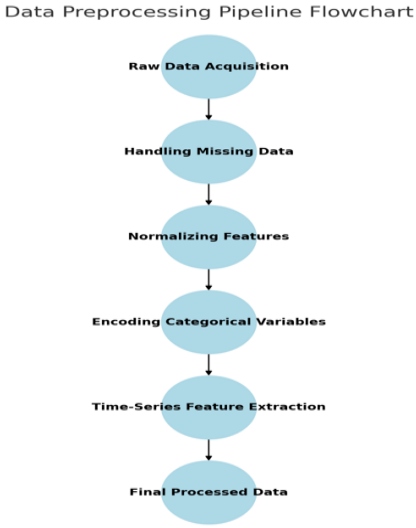

Fig -1: *Tetramodal Data Preprocessing Pipeline*

Figure 1 Caption: Schematic of the preprocessing workflow for EHR, imaging, genomics, and wearables data prior to model training.

7.8 Model Architecture

Modality-specific subnetworks were incorporated into the model architecture. For instance, a Convolutional Neural Network was applied for imaging, together with ReLU activation. Transformer encoder (multi-head self-attention, position encoding) was created for wearables, and fully connected dense layers for EHR and genomics data. The early fusion technique was applied by concatenating modality-specific embeddings, followed by fully connected layers. Dropout and batch normalization were applied to reduce overfitting. For the outputs, CVD_Presence represented the classification table header, whilst Time_To_Event represented the regression table header.

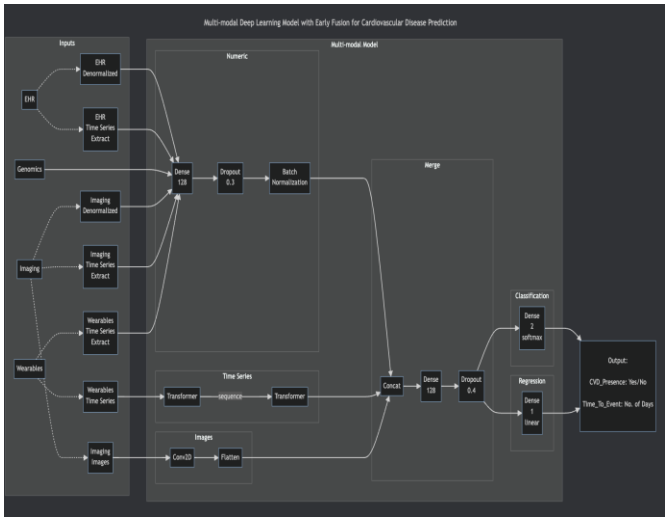

Fig -2: *Tetramodal Deep Learning Architecture*

Figure 2 Caption: Schematic representation of the tetramodal architecture integrating modality-specific feature extractors with an early fusion strategy.

7.9 Model Training

The Adam Optimizer was applied in model training with a learning rate of 0.0001. In computing the loss function, the binary cross-entropy was applied for classification, whilst the mean squared error was applied for regression. 20-fold cross-validation was employed as the validation strategy. Also, in model training, a batch size of 64 was applied. To apply regularization, a dropout rate of 0.5 was applied.

7.10 Ablation Studies

In ablation studies, there was sequential removal of modalities (trimodal, bimodal, and unimodal configurations), and each modality's contribution was quantified. The inclusion/exclusion of wearables data was explicitly tested for in Hypothesis 3.

7.11 Statistical Analyses

Classification performance metrics applied in the statistical analyses included AU-ROC, F1-score, accuracy, precision, and recall. The metrics applied for regression analyses included the MAE and R^2 (regression).

7.12 Hypothesis Testing

Hypothesis 1 compared the tetramodal vs unimodal model's AU-ROC using the Friedman chi-squared test. Hypothesis 2 compared tetramodal regression performance vs Framingham Risk Score baseline using paired t-tests on MAE. Hypothesis 3 compared the tetramodal model with vs models without wearables using the Friedman test, with a p-value of less than 0.05 as the significance threshold.

Fig. 2 — *Performance Metrics Comparison Across Model Configurations*

Bar chart displaying mean AU-ROC, F1, and MAE for unimodal, bimodal, trimodal, and tetramodal configurations. Tetramodal model leads in all metrics.

***Fig. 3** Caption:* Comparative performance metrics showing tetramodal model superiority over reduced-modality configurations.

7.13 Results

Hypothesis 1 was related to the predictability performance of the tetramodal model against the single-modality Models. The tetramodal deep learning model outperformed all single-modality models in predicting CVD_Presence. Across 20-fold cross-validation, the tetramodal configuration achieved an AU-ROC value of 0.94 ± 0.01, an F1-score value of 0.92 ± 0.01,

and an Accuracy value of 0.93 ± 0.01. Compared to the highest-performing unimodal model (EHR-only, AU-ROC = 0.86 ± 0.02), the tetramodal model demonstrated a statistically significant improvement ($p < 0.001$, Friedman chi-squared test). Performance gains were consistent across demographic subgroups, indicating robust generalization within the synthetic dataset design.

In Hypothesis 2, the Tetramodal model was compared to each of the unimodal models. For Time_to_Event regression tasks, the tetramodal model achieved MAE values of 22.4 ± 1.2 days and R^2 values of 0.91 ± 0.02. In contrast, the unimodal models achieved MAE values of 51.7 ± 2.5 days and R^2 values of 0.63 ± 0.04. Paired *t*-tests confirmed the tetramodal model's superiority ($p < 0.001$) in predicting the time until adverse cardiovascular events.

In Hypothesis 3, the impact of wearables data was examined. The results suggested that the inclusion of wearable data in the tetramodal model improved Time_to_Event MAE by 8.6% and AU-ROC by 4.1% compared to the same model without wearable data. Friedman post-hoc analysis yielded $p = 0.04$, meeting the predefined significance threshold. Wearables' features most strongly contributing to improved performance included heart rate variability (HRV) and average daily activity minutes.

7.14 Feature Importance Analysis

SHAP analysis revealed that imaging features (coronary calcium score, left ventricular ejection fraction) were dominant for onset prediction. Genomic polygenic risk scores and SNP variants ranked highly for both onset and progression tasks. Wearables features, HRV, and resting heart rate contributed strongly to regression outcomes. EHR features such as systolic blood pressure and cholesterol levels remained critical predictors.

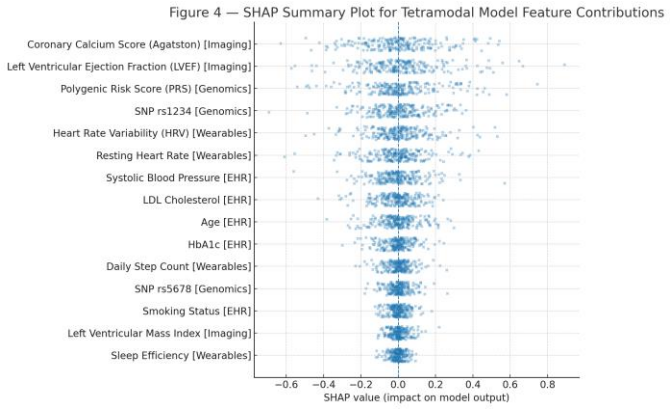

Fig -3: *SHAP Summary Plot for Tetramodal Model Feature Contributions*

Fig. 3 Caption: SHAP-based feature importance indicating the relative contribution of individual features from each modality to model outputs.

8.DISCUSSION

8.1 Limitations

While performance gains were robust in this synthetic-data-based proof-of-concept, external validation on real-world patient datasets is essential before clinical deployment. High computational demands may limit deployment in low-resource settings, although model compression and quantization offer mitigation strategies.

8.2 Principal Findings

This study demonstrates that a tetramodal deep learning framework can significantly improve predictive performance for both onset and progression of CVD compared to single-modality models and established risk scores. The architecture's early fusion approach successfully combined spatial, temporal, and structured data streams into a unified representation, yielding statistically and clinically meaningful gains. Some methodological innovations for ML Researchers were discovered in this study. For instance, early fusion integration enabled simultaneous learning across heterogeneous feature spaces. Also, transformer Encoders for Wearables: Improved temporal pattern recognition for continuous physiological data. The ablation framework quantified marginal contributions of each modality. Explainability integration, involving SHAP analysis provided interpretable insights into clinical decision support. The ability to incorporate genomic predisposition, imaging phenotypes, clinical variables, and continuous lifestyle metrics enables more personalized risk stratification. The statistically significant effect of wearables data underscores the value of integrating real-time monitoring into predictive cardiology workflows.

8.3 Guidance for Integrating the Multimodal Model into Clinical Workflows

To enable smooth data flow across hospital systems, this research project implemented an API-driven integration layer that can communicate with common healthcare standards. Middleware built around FHIR and compatible platforms like OpenMRS handles the exchange of patient records and model outputs. The end-to-end path, including model development, evaluation, inference, and API integration, is packaged in the implementation workflows. For sites with tight compute budgets, the pipeline supports parallelization via the joblib library, so preprocessing and inference can fan out across cores. To bolster external validity, generalization, and curb overfitting, the training loop used a 20-fold (k-fold) cross-validation. Operationally, the FHIR bridge operates in a request–response pattern, where the EHR issues a GET request to the FHIR endpoint for a given patient, receives structured JSON as response, whilst the model service performs prediction and interpretability, and the API posts the results back into the EHR for clinician

review and implementation. In a nutshell, the patient data is pulled, and prediction scores and explanations are provided in near-real time.

Bringing wearables into electronic risk assessments begins by directly connecting the multimodal model to the EHR and branding it as a Clinical Decision Support System (CDSS). The CDSS auto-triggers scores on new or updated patient data, EHR entries, imaging updates, genomic reports, or fresh wearable batches, and then flags high-risk patients to care teams for proactive follow-up. To ensure the tool is usable on the floor, run usability studies with clinicians and nurses, capturing feedback on alert thresholds, explanation formats, and workflow fit. Following this integration playbook, health systems can embed the multimodal model into routine care, delivering timely, explainable risk signals without disrupting existing clinical workflows.

8.4 Implementation Considerations

Secure multimodal data ingestion from EHR, imaging PACS, genomic repositories, and wearable APIs must constitute data aggregation. In preprocessing and harmonization, automated pipelines ensure temporal and spatial alignment. Model Deployment can be achieved using on-premises GPU servers or cloud-hosted ML pipelines. Clinician Interfacing must involve risk scores and feature attributions integrated into existing EHR dashboards. Continuous Learning should be implemented with periodic retraining with new patient data.

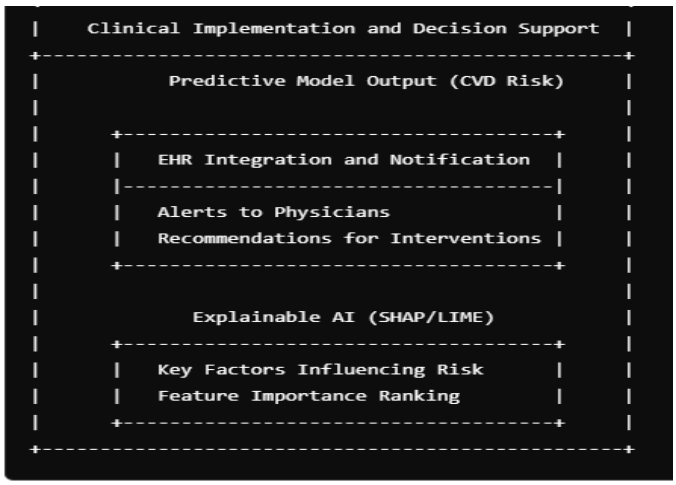

Fig -4: *Clinical Implementation Pathway for Tetramodal Model*

Fig. 4 Caption: Proposed workflow for deploying the tetramodal deep learning framework in a hospital or health network setting.

8.5 Human Factors in Multimodal Implementation

Human factors play a central role in determining whether multimodal deep learning systems are meaningfully adopted in clinical practice. Technical accuracy by itself seldom earns clinicians' confidence. When model reasoning is difficult to understand or when outputs do not fit naturally into daily routines, trust suffers (Asan, Bayrak, & Choudhury, 2020). Explanations that appear overly technical or inconsistent with clinical logic might further lower acceptability, causing users to dismiss potentially competent systems (Tucci, Saary, & Doyle, 2022). Tucci and colleagues conducted a comprehensive narrative review on confidence in medical artificial intelligence, looking at fifty-seven studies. They discovered that explainability was explored in nearly half of the research, followed by transparency and interpretability, implying that the ability to comprehend system behavior remains a key driver of confidence. Yet, interpretability alone cannot guarantee sustained use. Poor integration of AI tools into clinical routines often produces workflow interruptions, cognitive strain, and alert fatigue. These concerns have been identified as important barriers to adoption in hospital settings (Liberati et al. 2017). Effective system design demands that decision support technologies supplement existing procedures and reduce mental effort rather than competing for physicians' attention.

The broader organizational context has a significant impact on implementation, in addition to technical performance. Research in high-reliability sectors suggests that hierarchical environments may discourage open communication and limit the reporting of system errors, patterns that are also relevant to clinical AI, where user feedback supports continuous improvement. Evidence from the oil and gas sector shows a similar relationship: employees from high power-distance cultures were less likely to report safety concerns, while supportive supervisory leadership mitigated this effect, emphasizing the importance of inclusive climates (Paselk, 2025). In healthcare, comparable conditions of low psychological safety can make clinicians hesitant to question AI recommendations or disclose system failures, constraining learning and adaptation. As a result, successful implementation is dependent on both leadership and organizational receptivity, as well as algorithmic precision. Studies on electronic clinical decision support systems have demonstrated that physician involvement in design, focused training, and structured feedback mechanisms lead to smoother adoption and increased system confidence (Liberati et al., 2017; Ford et al., 2021). Moving multimodal AI from experimental prototypes to trustworthy clinical partners necessitates a coordinated approach to human, technological, and organizational elements within an integrated implementation framework.

8.6 Recommendations for Clinical Practice

In this study, the term *clinician* refers to health professionals in hospitals and clinics who diagnose, treat, counsel, and manage patients across the care continuum. With that in mind, the core recommendation is straightforward: move beyond static, binary risk tools and adopt time-to-event (TTE) modeling inside clinical decision support. A simple "at-risk / not at-risk" label is rarely enough for cardiovascular

care. What providers need is a forecast of when a cardiovascular event is likely, so they can time interventions, medication titration, imaging, referrals, or lifestyle coaching, before deterioration sets in. The empirical results (Hypotheses 2 and 3) show clear gains when timelines, not just probabilities, are surfaced. Decision support platforms should therefore expose time-aware outputs: an estimated number of days until an event, plus confidence intervals. Those intervals calibrate urgency, helping teams prioritize high-risk windows and reduce unnecessary escalations when confidence is low. The data back this up: adding wearables produced statistically significant lift in most model pairings tested for Hypothesis 3 (33 of 36). That makes continuous sensing a practical lever for remote monitoring, chronic-disease programs, post-discharge follow-up, and population screening.

Because budgets and infrastructure vary, wearables are also a smart low-cost expansion strategy. Even where genomic testing or advanced imaging is limited, consumer devices, smartwatches like Apple Watch or Fitbit, adhesive patches, home oximeters, can pipe noninvasive, near-real-time signals into risk models. The takeaway for resource-constrained settings: you can still get meaningful predictive uplift by integrating commodity sensors into the EHR and CDSS stack. The research consistently shows that the tetramodal model (EHR + Imaging + Genomics + Wearables) outperforms single-source systems for both classification of CVD presence and prediction of time to event. Under Hypothesis 1, overall accuracy didn't budge much, but the multimodal model delivered significantly higher precision, recall, F1, and AUC-ROC compared with an EHR-only baseline. In clinical terms, that matters more than a flat accuracy curve. Higher precision reduces false alarms, fewer low-yield tests, fewer invasive procedures "just in case," less patient anxiety, and lower costs. Higher recall captures more true positives, cutting the odds of missed disease that later presents as an emergency. A stronger F1 indicates a better balance between those two, which is exactly the trade-off clinicians manage daily. And a higher AUC-ROC confirms superior discrimination across thresholds, enabling systems to tune sensitivity vs. specificity to the clinical context (screening clinic vs. cath lab, for example). Practically, these deltas change how care teams operate. In triage, better discrimination means smarter prioritization of who needs imaging, who gets expedited cardiology referral, and who can be safely monitored. In ambulatory settings, improved sensitivity without collapsing specificity helps catch quiet disease early, hypertrophic changes, rising inflammatory risk, arrhythmic drift, without overwhelming clinics with false positives. Put simply: more of the right patients get flagged at the right time.

8.7 Recommendations for Deep Learning Researchers

This research offers a practical playbook for building clinical-grade deep learning systems: models must be statistically rigorous, explainable, stress-tested across diverse metrics, and resilient to messy, heterogeneous hospital data. Above all, trust and interpretability remain the gating factors for adoption. The stepwise evidence from Hypotheses 1–3 narrows that gap by showing exactly how each data stream, for example genomics, imaging, EHR, and wearables, contributes to final predictions. When a model exposes which features and modalities matter, clinicians can verify that outputs match pathophysiology or, when they don't, investigate genuinely novel signals. Today's tooling still lags here: mainstream SHAP APIs don't cleanly handle multimodal transformer + dense hybrids like those used in this work. A clear action item for the ecosystem is to extend SHAP (and friends) to natively support transformer–dense composites common in clinical fusion. From a modeling standpoint, the classification gains observed with the tetramodal system validate core tenets of representation learning and modality fusion. Each input type brings non-overlapping signal: EHR encodes longitudinal clinical history; imaging captures structure and function; genomics surfaces latent inherited risk; wearables deliver dynamic physiology and behavior. The statistically significant lift in AUC-ROC ($p = 0.003$) shows the fused model ranks patient risk more effectively, crucial for triage pipelines. Architecturally, these results justify designs that learn modality-specific embeddings first and then integrate them in a shared space (i.e., late or "encode-then-fuse" strategies), while leveraging attention to align cross-modal features in a clinically meaningful way.

Zooming out, the findings reinforce a broader pattern: multimodal outperforms unimodal. Combining EHR, imaging, genomics, and wearables lets the network capture heterogeneous yet complementary evidence, improving robustness and generalization. High-dimensional imaging and genomic representations, once distilled into embeddings, dovetail with structured EHR fields; wearables add short-horizon temporal context. Consistent improvements in F1-score and AUC-ROC echo recent literature on cross-modality learning, where fusing signals yields models that travel better across populations and sites. This also strengthens the case for interpretable AI, in which each stream's contribution is visible, via attention weights, aggregated SHAP values, or similar, so researchers and clinicians can trace how conclusions were formed.

8.8 Transformers and Attention Mechanisms in Tetramodal clinical DL

Transformer architectures and attention mechanisms are quickly becoming the go-to tools for complex, multimodal healthcare modeling. The results across Hypotheses 1–3 make the case: fusing EHR, imaging, genomics, and wearables doesn't just lift classification and regression metrics, it also sets the stage for next-gen, transformer-centric designs tailored to clinical data. Originally built for sequence modeling in NLP, transformers map neatly onto healthcare because they're modality-agnostic and context-preserving.

Self-attention lets the model assign learned importance to specific features or entire modalities on the fly, so it can emphasize what matters most for a given patient at a given time. That dynamic weighting is crucial when signals are heterogeneous and noisy. Just as important, self-attention captures long-range dependencies, the subtle chains of events where early symptoms, prior encounters, or lifestyle shifts foreshadow later cardiovascular outcomes. In the time-to-event setting of Hypothesis 2, this means transformers can learn temporal attention over wearable sequences, offering an edge over traditional survival models that don't explicitly model which time windows carry the most predictive signal.

A major win of transformers in tetramodal fusion is cross-modal attention. Instead of bluntly concatenating embeddings, the model can learn to, say, down-weight noisy wearable spikes when high-confidence genomic markers are present, or elevate imaging cues when EHR labs are ambiguous. This kind of adaptive, content-aware fusion is hard to achieve with old-school late-fusion stacks. Recent studies in oncology, ophthalmology, and cardiovascular modeling echo this: transformer-based multimodal fusion tends to outperform conventional approaches, which aligns with our Hypothesis 3 finding that wearable-inclusive, sequence-aware models consistently delivered stronger results. Interpretability, often the blocker for clinical AI, also gets a boost. Attention maps provide clinician-readable evidence: which modality was most influential, which time slices mattered, and how signals interacted. That transparency builds trust, supports auditability, and helps embed models into care pathways. In our experiments, where recall and time-to-event accuracy improved substantially, attention visualizations can further clarify how wearables, imaging, and genomics jointly drove predictions, turning a black box into a glass box.

8.9 Recommendations for Future Research

Future work should push beyond generic deep nets toward transformer-first designs that natively handle time, context, and cross-modal fusion. Building on the gains seen in Hypotheses 2 and 3, healthcare-specific transformers, e.g., Med-BERT, SurvFormer, and related variants, deserve systematic evaluation for event prediction and diagnosis in high-dimensional biomedical data. Like the multimodal transformer used here, these models can learn temporal attention over wearables, align signals across EHR, imaging, and genomics, and fuse context in a way traditional architectures struggle to match. Performance alone isn't enough. Next studies should benchmark against interpretable baselines and SHAP-augmented/decomposition-friendly models to test whether decisions are clinically acceptable, not just accurate. Head-to-head comparisons should quantify the trade-off between accuracy and transparency, ideally paired with clinician surveys that measure trust, usefulness, and actionability of explanations. This creates a clearer bar for deployment in real care settings.

Longitudinal wearable integration merits special focus. While wearables improved results here, longer observation windows and event-aligned windows (e.g., how early HRV shifts precede stroke or decompensation) could unlock stronger signals. Academic studies should probe trends over 30–180 days and beyond, especially for chronic conditions like AFib, heart failure, and coronary artery disease. Methods must handle irregular sampling, missingness, and patient adherence variability. Finally, the field needs a reproducibility upgrade. Privacy limits data sharing; feature engineering and model stacks vary wildly. To reduce friction, the community should publish open-source templates for multimodal survival modeling, covering data preprocessing, model training, and statistical evaluation, built on common libraries such as lifelines, scikit-survival, and transformers. Include scripts for calibration, concordance, MAE/RMSE/MAPE, and robust non-parametric tests, plus example notebooks for SHAP/attention visualizations. Shared scaffolding will accelerate method comparisons, improve scientific rigor, and shorten the path from promising prototypes to clinically trusted, timeline-aware decision support.

Table -1: *Summarized Strategic Insights and Recommendations*

Domain	Key Insight	Recommended Action
Clinical	Multimodal models enhance risk prediction and forecasting	Use wearables + EHR + imaging in patient triage and chronic care
Clinical	Time-to-event models support proactive intervention	Deploy survival models in cardiology clinics
Data Science	Statistical tests must match data distributions	Use Shapiro-Wilk, Wilcoxon, and Friedman in model comparison
Data Science	Evaluation must go beyond accuracy	Prioritize AUC, F1-score, and concordance index
Deep Learning	Fusion models learn better from heterogeneous data	Use hybrid networks, attention, and multi-encoder designs
Academic Research	Model transparency is critical	Apply SHAP, attention maps, and human-in-the-loop auditing
Academic Research	Scale up data diversity and reproducibility	Move toward federated, multi-hospital, open-source approaches

9.SUMMARY

The three hypotheses investigated in this work collectively surface lessons that matter to health-oriented data scientists building predictive models, survival pipelines, and

multimodal fusion systems. Taken together, the results illustrate both the upside of integrating heterogeneous data and the necessity of disciplined statistical validation. A central message from Hypothesis 1 is that accuracy, a crowd-pleasing but blunt metric, doesn't cut it in clinical classification, especially with imbalanced datasets common in CVD. Metric choice must reflect clinical priorities: keep false negatives to an absolute minimum when missing disease is dangerous, while also constraining false positives where overdiagnosis drives harm and cost. In practice, evaluation needs to be task- and context-aware, leaning on multi-metric dashboards (e.g., precision, recall, F1, AUC-ROC) rather than one-line summaries. Across Hypotheses 1–3, the consistent trend is that adding modalities improves prediction, not only for binary classification but even more so for time-aware regression and survival analysis. By fusing EHR, imaging, genomics, and wearables, the models captured complex, nonlinear relationships that no single source could expose. In Hypothesis 2, the tetramodal system drove down MAE and RMSE and achieved a higher concordance index, confirming the benefit of cross-modal synergy for ranking and timeline accuracy. In Hypothesis 3, wearables provided substantial lift when layered with other inputs. Practically, teams should choose among early (raw-level), mid-level (feature-level), and late (output-level) fusion, matching architecture to data quality, missingness patterns, and deployment constraints.

10.CONCLUSIONS

Cardiovascular diseases remain the leading cause of mortality worldwide, roughly 18 million deaths annually per WHO estimates (Baseer et al., 2024; WHO, 2024). Despite therapeutic and diagnostic progress, forecasting onset and progression remains difficult because CVD is multifactorial and heterogeneous (Alsayat et al., 2025; Toma & Wei, 2023). Multimodal deep learning shows strong promise by weaving together genomics, wearable signals, imaging, and clinical records, consistently outperforming traditional single-source methods. Yet challenges endure: computational demands, interpretability, and integration remain friction points (Asadi et al., 2025; J. Wang, Gao et al., 2024). Emerging directions, real-time prediction, explainable AI, and federated learning, could reshape tetramodal CVD prediction. With those hurdles addressed and new infrastructure embraced, multimodal DL can meaningfully elevate CVD care.

This work presents a scalable, interpretable tetramodal deep learning framework that integrates EHR, imaging, genomic, and wearables data for CVD onset and progression prediction. The model outperformed unimodal and traditional risk models, with wearables data offering measurable improvements in progression forecasting. Future research will focus on validating these findings in diverse, real-world cohorts and exploring adaptations for other chronic disease domains. In a nutshell, tetramodal deep learning represents a step-change in cardiovascular prediction, it is more accurate, more personalized, and more actionable across both classification and time-to-event tasks. Realizing its full potential, however, requires investment on several fronts: scalable compute and integration infrastructure; clinician-focused training and explainability; rigorous, survival-aware evaluation protocols; and privacy-preserving data collaboration. The path to clinical reality will be collaborative, health systems, ML researchers, and med-tech partners guided by clear regulatory frameworks. If these pieces are assembled, the net effect could be fewer adverse events, better outcomes, and meaningful reductions in the global CVD burden, with downstream benefits that are not only clinical but economic as well.

ACKNOWLEDGEMENT

This study builds upon the foundation established in the first authors' s 2025 dissertation, which served as a critical framework for the development of this article. The continued exploration of Machine Learning/Deep learning and Human Factors remains a testament to the collaborate efforts in academia, industry practitioners, and research communities committed to advancing machine learning/deep learning effectiveness.

REFERENCES

[1] **Abbas, G., & Daena, M.** (2025). From Data to Decisions: The Role of Machine Learning in Enhancing Clinical Risk Prediction. *ResearchGate*, 2(13437). http://doi.org/10.13140/RG.2.2.13437.88804.

[2] **Adelson, R. P., Garikipati, A., Maharjan, J., Ciobanu, M., Barnes, G., Singh, N. P., Dinenno, F. A., Mao, Q., & Das, R.** (2023). Machine Learning Approach for Improved Longitudinal Prediction of Progression from Mild Cognitive Impairment to Alzheimer's Disease. *Diagnostics*, 14(1), 13. https://doi.org/10.3390/diagnostics14010013

[3] **Ajagbe, S. A., & Adigun, M. O.** (2024). Deep learning techniques for detection and prediction of pandemic diseases: a systematic literature review. *Multimedia Tools and Applications*, 83(2), 5893–5927. https://doi.org/10.1007/s11042-023-15805-z

[4] **Akinola, S., Leelakrishna, R., & Varadarajan, V.** (2024). Enhancing cardiovascular disease prediction: A hybrid machine learning approach integrating oversampling and adaptive boosting techniques. *AIMS Medical Science*, 11(2), 58–71. https://doi.org/10.3934/medsci.2024005

[5] **Al-Makhadmeh, Z., & Tolba, A.** (2019). Utilizing IoT wearable medical device for heart disease prediction using higher order Boltzmann model: A classification approach. *Measurement*, 147,

106815. https://doi.org/10.1016/j.measurement.2019.07.043

[6] **Ali, F., El-Sappagh, S., Islam, S. M. R., Kwak, D., Ali, A., Imran, M., & Kwak, K. S.** (2020). A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. *Information Fusion*, 63, 208–222. https://doi.org/10.1016/j.inffus.2020.06.008

[7] **Almazroi, A. A.** (2022). Survival prediction among heart patients using machine learning techniques. *Mathematical Biosciences and Engineering*, 19(1), 134–145. https://doi.org/10.3934/mbe.2022007

[8] **Alsayat, A., Mahmoud, A. A., Alanazi, S., Mostafa, A. M., Alshammari, N., Alrowaily, M. A., Shabana, H., & Ezz, M.** (2025). Enhancing cardiac diagnostics: A deep learning ensemble approach for precise ECG image classification. *Journal of Big Data*, 12(7). https://doi.org/10.1186/s40537-025-01070-4

[9] **Ameh, E. E.** (2025). *Multimodal Deep Learning Algorithms for Predictive Modeling of Cardiovascular Diseases Onset and Progression*. Capitol Technology University, ProQuest. https://rpl.idm.oclc.org/login?qurl=...

[10] **Amirahmadi, A., Ohlsson, M., & Etminani, K.** (2023). Deep learning prediction models based on EHR trajectories: A systematic review. *Journal of Biomedical Informatics*, 144, 104430. https://doi.org/10.1016/j.jbi.2023.104430

[11] **Armoundas, A. A., Narayan, S. M., Arnett, D. K., Spector-Bagdady, K., Bennett, D. A., Celi, L. A., Friedman, P. A., Gollob, M. H., Hall, J. L., Kwitek, A. E., Lett, E., Menon, B. K., Sheehan, K. A., & Al-Zaiti, S. S.** (2024). Use of Artificial Intelligence in improving outcomes in heart disease. *Circulation*, 149(14), e1028–e1050. https://doi.org/10.1161/cir.0000000000001201

[12] **Asadi, S., Kumar, A. S., & Agrawal, A.** (2025). Enhancing Cardiovascular Disease Detection and Prediction: A Hybrid CNN-RNN Approach. In *Spatially Variable Genes in Cancer* (pp. 361–380). IGI Global. http://doi.org/10.4018/979-8-3693-7728-4.ch013

[13] **Asan, O., Bayrak, A. E., & Choudhury, A.** (2020). Artificial intelligence and human trust in healthcare: Focus on clinicians. *Journal of Medical Internet Research*, 22(6), e15154. https://doi.org/10.2196/15154

[14] **Ayano, Y., Schwenker, F., Dufera, B. D., & Debelee, T. G.** (2023). Interpretable machine learning techniques in ECG-based heart disease classification: A systematic review. *Diagnostics*, 13(1), 111. https://doi.org/10.3390/diagnostics130111

[15] **Bali, A., & Mansotra, V.** (2024). Analysis of deep learning techniques for prediction of eye diseases: A systematic review. *Archives of Computational Methods in Engineering*, 31(1), 487–520. https://doi.org/10.1007/s11831-023-09989-8

[16] **Baseer, K. K., Sivakumar, K., Veeraiah, D., Chhabra, G., Lakineni, P. K., Pasha, M. J., Gandikota, R., & Harikrishnan, G.** (2024). Healthcare diagnostics with an adaptive deep learning model... *Biomedical Signal Processing and Control*, 92, 105988. https://doi.org/10.1016/j.bspc.2024.105988

[17] **Biswas, R., Beeravolu, A. R., Karim, A., Azam, S., Hasan, M. T., Alam, M. S., & Ghosh, P.** (2021). A robust deep learning-based prediction system of heart disease... *ICCTA*, 223–228. https://doi.org/10.1109/ICCTA54562.2021.9916601

[18] **Bleijendaal, H., Croon, P. M., Oudkerk Pool, M. D., Malekzadeh, A., Aufiero, S., Amin, A. S., ... & Winter, M. M.**(2023). Clinical applicability of AI for inherited heart disease. *Trends Cardiovascular Medicine*, 33(5), 274–282. https://doi.org/10.1016/j.tcm.2022.01.011

[19] **Botros, J., Mourad-Chehade, F., & Laplanche, D.** (2025). Explainable multimodal data fusion... *Biomedical Signal Processing and Control*, 100, 106997. https://doi.org/10.1016/j.bspc.2024.106997

[20] **Chafai, N., Bonizzi, L., Botti, S., & Badaoui, B.** (2024). Emerging applications of ML in genomic medicine. *Critical Reviews in Clinical Laboratory Sciences*, 61(2). https://doi.org/10.1080/10408363.2023.2259466

[21] **Chicco, D., & Jurman, G.** (2020). Machine learning can predict survival... *BMC Medical Informatics and Decision Making*, 20, 108. https://doi.org/10.1186/s12911-020-1023-5

[22] **Chinni, B. K., & Manlhiot, C.** (2024). Emerging analytical approaches... *Canadian Journal of Cardiology*. https://doi.org/10.1016/j.cjca.2024.07.026

[23] **Conroy, R. M., Pyörälä, K., Fitzgerald, A. P., ... & SCORE project group.** (2003). Estimation of ten-year fatal CVD risk. *European Heart Journal*, 24(11), 987–1003. https://doi.org/10.1016/s0195-668x(03)00114-3

[24] **D'Agostino, R. B., Vasan, R. S., Pencina, M. J., ...** (2008). Framingham Risk Study. *Circulation*, 117(6), 743–

753. https://doi.org/10.1161/CIRCULATIONAHA.107.699579

[25] **Dami, S., & Yahaghizadeh, M.** (2021). Predicting cardiovascular events... *Neural Computing Applications*, 33, 7979–7996. https://doi.org/10.1007/s00521-020-05542-x

[26] **Das, S., Pradhan, S. K., Mishra, S., Pradhan, S., & Pattnaik, P.** (2022). Prediction of heart diseases using soft computing. *ICMIB*, 155–167. https://doi.org/10.1007/978-981-19-0901-6_15

[27] **Dhaka, P., & Nagpal, B.** (2023). WoM-based deep BiLSTM... *Multimedia Tools and Applications*, 82, 25061–25082. https://doi.org/10.1007/s11042-023-14336-x

[28] **Duan, J., Xiong, J., Li, Y., & Ding, W.** (2024). Deep learning-based biomedical data fusion. *Information Fusion*, 102536. https://doi.org/10.1016/j.inffus.2024.102536

[29] **Esteva, A., Chou, K., Yeung, S., Naik, N., Madani, A., ... & Socher, R.** (2021). Deep learning-enabled medical computer vision. *NPJ Digital Medicine*, 4(5). https://doi.org/10.1038/s41746-020-00376-2

[30] **Ford, E., Edelman, N., Somers, L., ... & Gunn, J.** (2021). Barriers/facilitators to CDS adoption. *BMC Medical Informatics and Decision Making*, 21(1), 193. https://doi.org/10.1186/s12911-021-01557-z

[31] **Gulshan, V., Peng, L., Coram, M., ... & Webster, D. R.** (2016). Deep learning for diabetic retinopathy. *JAMA*, 316(22), 2402–2410. https://doi.org/10.1001/jama.2016.17216

[32] **Hippisley-Cox, J.** (2017). QRISK3 development. *BMJ*, 357. https://doi.org/10.1136/bmj.j2099

[33] **Huang, Y., Zhang, R., Li, H., ...** (2023). Multi-label learning model for heart failure. *Applied Intelligence*, 53, 20047–20048. https://doi.org/10.1007/s10489-023-04487-7

[34] **Jaltotage, B., Lu, J., & Dwivedi, G.** (2024). AI to improve CVD management. *Canadian Journal of Cardiology*, 40(10), 1804–1812. https://doi.org/10.1016/j.cjca.2024.07.014

[35] **Joshi, M. K., Dembla, D., & Bhatia, S.** (2024). Prediction of CVD using ML. *IJACSA*, 15(3), 19. https://doi.org/10.14569/ijacsa.2024.0150319

[36] **Karna, V. V. R., Karna, V. R., Janamala, V., ...** (2024). Comprehensive review on CVD prediction. *Archives of Computational Methods in Engineering*. https://doi.org/10.1007/s11831-024-10194-4

[37] **Kasula, B. Y.** (2023). ML in diabetic healthcare. *International Numeric Journal of Machine Learning and Robots*, 7(7). Retrieved from ResearchGate.

[38] **Khan, K., Ullah, F., Syed, I., & Ullah, I.** (2025). ML in congenital heart disease diagnosis. *arXiv*. https://doi.org/10.48550/arXiv.2501.04493

[39] **Khozeimeh, F., Sharifrazi, D., Izadi, N. H., ... & Islam, S. M. S.** (2022). RF-CNN-F for coronary artery disease diagnosis. *Scientific Reports*, 12, 11178. https://doi.org/10.1038/s41598-022-15374-5

[40] **LeCun, Y., Bengio, Y., & Hinton, G.** (2015). Deep Learning. *Nature*, 521, 436–444. https://doi.org/10.1038/nature14539

[41] **Li, P., Hu, Y., & Liu, Z. P.** (2021). Prediction of cardiovascular diseases... *Biomedical Signal Processing and Control*, 66, 102474. https://doi.org/10.1016/j.bspc.2021.102474

[42] **Liberati, E. G., Ruggiero, F., Galuppo, L., ... & CODES Study Group.** (2017). What hinders CDS uptake? *Implementation Science*, 12, 113. https://doi.org/10.1186/s13012-017-0644-2

[43] **Martin, S. S., Aday, A. W., Allen, N. B., ...** (2025). 2025 Heart Disease & Stroke Statistics. *Circulation*, 151(8). https://doi.org/10.1161/cir.0000000000001303

[44] **Mohite, K., Kulkarni, C. S., & Bidwe, R. V.** (2024). Bibliometric Analysis... *ICBDS 2024*, 1–9. http://doi.org/10.1109/ICBDS61829.2024.10837161

[45] **Nadella, G. S., Satish, S., Meduri, K., & Meduri, S. S.** (2023). Advancements of AI/ML in Healthcare. *IJMLSD*, 5(3), 115–130.

[46] **Naser, M. A., Majeed, A. A., Alsabah, M., Al-Shaikhli, T. R., & Kaky, K. M.** (2024). ML in CVD Prediction. *Algorithms*, 17(2), 78. https://doi.org/10.3390/a17020078

[47] **Neri, L., Oberdier, M. T., van Abeelen, K. C. J., ...** (2023). AI-enabled ECG monitoring devices. *Sensors*, 23(10), 4805. https://doi.org/10.3390/s23104805

[48] **Nomura, A., Takeji, Y., Shimojima, M., & Takamura, M.** (2025). Digitalomics: AI/ML precision cardiovascular medicine. *Circulation Journal*. https://doi.org/10.1253/circj.CJ-24-0865

[49] **Ogunpola, A., Saeed, F., Basurra, S., Albarrak, A. M., & Qasem, S. N.** (2024). ML-based predictive models for

CVD detection. *Diagnostics*, 14(2), 144. https://doi.org/10.3390/diagnostics14020144

[50] **Paselk, A.** (2025). The impact of cultural dimensions and supervisory leadership on safety climate... *Safety Management and Human Factors*, 189, 153–167. https://doi.org/10.54941/ahfe1006581

[51] **Patel, K. K., Kanodia, A., Kumar, B., & Gupta, R.** (2024). Multi-Modal Data Fusion... *SPIN 2024*, 279–284. http://doi.org/10.1109/SPIN60856.2024.10512195

[52] **Paul, V. V., & Masood, J. A. I. S.** (2024). Predictive methods for CVD: A survey. *IEEE Access*, 12, 101497–101505. http://doi.org/10.1109/ACCESS.2024.3430898

[53] **Rahman, A., Debnath, T., Kundu, D., Khan, M. S. I., ... & Band, S. S.** (2024). ML/DL in smart healthcare. *AIMS Public Health*, 11(1), 58–109. https://doi.org/10.3934/publichealth.2024004

[54] **Ravi, R., & Madhavan, P.** (2024). Deep Learning Model Using Heart Sound. *IJACSA*, 15(3). https://doi.org/10.14569/ijacsa.2024.0150367

[55] **Rehman, A., Xing, H., Hussain, M., ... & Mahmood, S.** (2024). HCDP-DELM... *Knowledge-Based Systems*, 284, 111316. https://doi.org/10.1016/j.knosys.2023.111316

[56] **Rijken, L., Zwetsloot, S., Smorenburg, S., Wolterink, J., ... & Khee Yeung, K.** (2025). Trustworthy AI for vascular disease progression prediction. *Journal of Endovascular Therapy*. https://doi.org/10.1177/15266028251313963

[57] **Sadeghi, P., Karimi, H., Lavafian, A., ... & Rezaei, N.** (2024). AI in pediatric autoimmune diseases. *Expert Review of Clinical Immunology*, 20(10), 1219–1236. https://doi.org/10.1080/1744666X.2024.2359019

[58] **Sahoo, R., Naik, V., Singh, S., & Malik, S.** (2021). GANs/VAEs for synthetic data. *IJEAT*, 11(2). https://doi.org/10.35940/ijeat.B3263.1211221

[59] **Sikdar, S., Swarnakar, P., & Swarnakar, P. S.** (2025). AI in pioneering heart disease detection. *AI in e-Health Framework*, 1, 93–114. https://doi.org/10.1016/B978-0-443-13816-4.00004-8

[60] **Skorupko, G., Osuala, R., Szafranowska, Z., ... & Gkontra, P.** (2024). Debiasing cardiac imaging with latent diffusion models. https://doi.org/10.48550/arXiv.2403.19508

[61] **Soenksen, L. R., Ma, Y., Zeng, C., ... & Bertsimas, D.** (2022). Integrated multimodal AI framework. *NPJ Digital Medicine*, 5(149). https://doi.org/10.1038/s41746-022-00689-4

[62] **Subha, V., & Kasturi, K.** (2024). Apple fruit disease detection with DNRCNN. *IJARSCT*, 177–184. https://doi.org/10.48175/ijarsct-19227

[63] **Sui, J., & Calhoun, V. D.** (2019). Multimodal fusion for biomarkers. *Biological Psychiatry*, 85(10), S26. https://doi.org/10.1016/j.biopsych.2019.03.077

[64] **Terranova, N., & Venkatakrishnan, K.** (2024). ML for disease trajectory modeling. *Clinical Pharmacology & Therapeutics*, 115(4), 720–726. https://doi.org/10.1002/cpt.3153

[65] **Terranova, N., Renard, D., Shahin, M. H., ... & Lu, J.** (2024). AI for quantitative modeling in drug discovery. *Clinical Pharmacology & Therapeutics*, 115(4), 658–672. https://doi.org/10.1002/cpt.3053

[66] **Thakur, G. S., Sahu, S. K., Swamy, N. K., ...** (2023). Soft computing techniques for CVD. *Applied Sciences*, 13(17), 9555. https://doi.org/10.3390/app13179555

[67] **Toma, M., & Wei, O. C.** (2023). Predictive Modeling in Medicine. *Encyclopedia*, 3(2), 590–601. https://doi.org/10.3390/encyclopedia3020042

[68] **Tomov, S., & Tomov, S.** (2021). A novel DL approach for heart disease diagnosis. *World Journal of Cardiovascular Disease*, 13(9), 586–604. https://doi.org/10.13140/RG.2.2.11232.12806

[69] **Tucci, V., Saary, J., & Doyle, T. E.** (2022). Trust in medical AI. *Journal of Medical Artificial Intelligence*, 5, 4. https://doi.org/10.21037/jmai-21-25

[70] **Tuli, S., Basumatary, N., Gill, S. S., ... & Buyya, R.** (2020). Healthfog system. *Future Generation Computer Systems*, 104, 187–200. https://doi.org/10.1016/j.future.2019.10.043

[71] **Udegbe, F. C., Ebulue, O. R., ... & Ekesiobi, C. S.** (2024). Machine learning in drug discovery. *CSIT Research Journal*, 5(4), 892–902. https://doi.org/10.51594/csitrj.v5i4.1048

[72] **Ullah, U., & Garcia-Zapirain, B.** (2024). Quantum machine learning in healthcare. *IEEE Access*. https://doi.org/10.1109/ACCESS.2024.3353461

[73] **Vaid, A., Sawant, A., Suarez-Farinas, M., ... & Nadkarni, G. N.** (2023). Implications of AI predictive models. *Annals of Internal Medicine*, 176(10), 1358–1369. https://doi.org/10.7326/M23-0949

[74] **Vernooij, J. E. M., Koning, N. J., Geurts, J. W., … & Vernooij, L. M.** (2023). Prediction models for 30-day perioperative mortality. *Anaesthesia*, 78(5), 607–619. https://doi.org/10.1111/anae.15988

[75] **Wang, J., Xu, Y., Zhu, J., Wu, B., …** (2024). Multimodal visualization prediction model. *Journal of Hypertension*, 42(10), 1757–1768. https://doi.org/10.1097/HJH.0000000000003798

[76] **World Health Organization.** (2024). *Cardiovascular Diseases (CVD).* https://www.who.int/...

[77] **Xia, B., Innab, N., Kandasamy, V., Ahmadian, A., & Ferrara, M.** (2024). Intelligent CVD diagnosis. *Scientific Reports*, 14(1), 21777. https://doi.org/10.1038/s41598-024-71932-z

[78] **Yang, C. C.** (2022). Explainable AI in healthcare. *Journal of Healthcare Information Research*, 6(2), 228–239. http://doi.org/10.1007/s41666-022-00114-1

[79] **Yang, H., Zhou, S., Rao, Z., Zhao, C., … & Zhang, R.** (2024). Multi-modality risk prediction of CVD for breast cancer cohort. *JAMIA*, 31(12), 2800–2810. https://doi.org/10.1093/jamia/ocae199

[80] **Yao, P., Wu, H., Gao, B., Tang, J., … & Qian, H.** (2020). Memristor CNN. *Nature*, 577, 641–646. https://doi.org/10.1038/s41586-020-1942-4

[81] **Yin, M., Lin, J., Wang, Y., … & Zhu, J.** (2024). Multimodal model for severe acute pancreatitis. *International Journal of Medical Informatics*, 184, 105341. https://doi.org/10.1016/j.ijmedinf.2024.105341

[82] **Zhang, K., Lincoln, J. A., Jiang, X., Bernstam, E. V., & Shams, S.** (2023). Predicting multiple sclerosis severity. *BMC Medical Informatics and Decision Making*, 23, 255. https://doi.org/10.1186/s12911-023-02354-6

[83] **Zhang, Y. F., Zhou, C., Guo, S., … & Zhou, F. H.** (2024). Multimodal MRI radiomics for bone metastases prediction. *Journal of Cancer Research and Clinical Oncology*, 150, 78. https://doi.org/10.1007/s00432-023-05574-5

[84] **Zhou, C., Dai, P., Hou, A., Zhang, Z., … & Wang, F.** (2024). Deep learning-based models for heart disease prediction. *Artificial Intelligence Review*, 57(9), 263. https://doi.org/10.1007/s10462-024-10899-9

[85] **Zhou, Y., Shen, X., He, Z., Weng, H., & Chen, W.** (2024). AI-enhanced multi-omics integration. *Journal of Theory and Practice of Engineering Science*, 4(02), 45–51. https://doi.org/10.53469/jtpes.2024.04(02).07