Sustainable, Equity-Focused AI Models for Reducing Menopause-Related Cardiovascular Burden in Low- and Middle-Income Countries

Hannah Imtiaz, Menaha Bhandari, Thanigaivelu P. S.
Department of Computing Technologies
SRM Institute of Science and Technology, Kattankulathur
Chennai, Tamilnadu, India
hi0642@srmist.edu.in, mb4283@srmist.edu.in, thanigap2@srmist.edu.in

Abstract

Cardiovascular disease (CVD) remains the leading cause of mortality among women worldwide, with risk increasing substantially following menopause due to hormonal and metabolic changes. Despite this well-documented clinical relationship, existing machine learning-based CVD prediction models rarely incorporate menopause-specific physiological transitions or socioeconomic disparities that disproportionately affect women in low- and middle-income countries (LMICs). This study proposes an equity-aware machine learning framework that integrates menopause-stage indicators, menopause–cardiovascular interaction features, and socioeconomic proxies with traditional cardiovascular risk factors. A Random Forest classifier was implemented using clinically grounded features derived from the NHANES 2017–2020 dataset. To address the severe class imbalance inherent in cardiovascular screening data, class-weight balancing and automated threshold optimization were applied. The optimized model achieved an accuracy of 89.2%, a recall of 0.875, and a ROC-AUC of 0.914, confirming the feasibility of menopause-aware, equity-focused AI for preventive cardiovascular screening in resource-constrained healthcare settings. Feature importance analysis further validated that menopause–cholesterol and menopause–blood pressure interaction terms were the strongest predictors of cardiovascular risk in this cohort.

Keywords: Cardiovascular disease, menopause, Random Forest, machine learning, health equity, LMIC, class imbalance, federated learning, fairness-aware AI, data quality.

How to cite this article: Imtiaz H, Bhandari M, Thanigaivelu PS. Sustainable, Equity-Focused AI Models for Reducing Menopause-Related Cardiovascular Burden in Low- and Middle-Income Countries. Int J Drug Deliv Technol. 2026;16(55s): 252-259. DOI: 10.25258/ijddt.16.55s.28

Introduction

disease (CVD) is the leading cause of death globally, and women face a disproportionately elevated risk following the menopausal transition. The decline in circulating estrogen levels during and after menopause triggers a cascade of metabolic and vascular changes—including dyslipidemia, endothelial dysfunction, increased arterial stiffness, and adipose redistribution—each of which independently elevates cardiovascular risk [1]. Despite this well-established clinical evidence, the majority of machine learning (ML)-based CVD prediction models are developed without explicit consideration of menopausal status as a risk modifier, effectively averaging risk patterns across heterogeneous female populations with distinct physiological profiles.

This oversight is compounded in low- and middle-income countries (LMICs), where postmenopausal women frequently encounter additional barriers including delayed diagnosis, fragmented health records, limited access to preventive care, and distinct cardiovascular risk factor distributions [2]. Current AI-based predictive systems are predominantly trained on datasets from high-income regions, limiting their generalizability and clinical relevance in LMIC healthcare environments. The resulting models may systematically underperform for precisely the populations that carry the highest unmet burden of cardiovascular disease.

A further challenge in developing effective CVD screening tools is the inherent class imbalance in population-level cardiovascular datasets. Confirmed CVD cases represent a small minority of screened individuals, and models optimized solely for overall accuracy tend to collapse into majority-class prediction, achieving high accuracy while failing to identify the high-risk minority—a critical failure mode in preventive healthcare contexts [3]. This challenge necessitates specialized modeling strategies including class-weight balancing, threshold optimization, and the use of sensitivity-focused evaluation metrics.

In parallel, concerns around bias, fairness, and equity in healthcare AI have gained increasing prominence. Narrative and systematic reviews demonstrate that standard performance metrics often obscure disparities across demographic subgroups, particularly when biological sex, life-stage variables, and socioeconomic attributes are not explicitly modeled [6]. Fairness-aware methodologies are rarely integrated into cardiovascular AI pipelines as core design principles and are typically applied only as post-hoc evaluations.

This work addresses these interconnected gaps by proposing a sustainable, equity-focused cardiovascular disease prediction framework that explicitly incorporates menopausal stage, menopause–risk interaction features, and LMIC-representative socioeconomic proxies. Rather than proposing entirely new algorithms, this study integrates established ML and federated learning techniques with structured data quality prioritization [2] and fairness-aware evaluation methodologies [6]. The proposed framework is designed with deployability in resource-constrained settings as a central objective.

1.1 Contributions

The primary contributions of this work are:

- A menopause-aware cardiovascular risk prediction framework incorporating clinically grounded age-based menopausal stage proxies and menopause–risk interaction features.
- Integration of socioeconomic indicators representing income level and healthcare access disparities characteristic of LMIC environments.
- Implementation of a class-balanced Random Forest classifier with automated threshold optimization, achieving high sensitivity under severe class imbalance.
- A data quality prioritization approach using AHP-informed sensitivity analysis [2] applied to the cardiovascular screening context.
- Subgroup-level fairness evaluation across menopause stages and socioeconomic strata, with analysis of feature importance for clinical interpretability.

2 Related Work

2.1 Machine Learning for Cardiovascular Disease Prediction

Supervised machine learning methods have been extensively applied to CVD prediction using clinical, demographic, and lifestyle data. El-Sofany [1] demonstrated that ensemble and tree-based classifiers consistently outperform linear models in capturing nonlinear interactions among cardiovascular risk factors including age, blood pressure, cholesterol, smoking behavior, and comorbidities. Paul and Masood [3] conducted a comprehensive survey of cardiovascular prediction methods, noting that no single algorithm consistently dominates across datasets, with predictive performance strongly influenced by data characteristics and preprocessing strategies. Sunilkumar and Kumaresan [7] reviewed deep learning and transfer learning applications in cardiology, finding that while deep models achieve performance gains in structured clinical data and ECG analysis, they remain dataset-specific and lack cross-population validation.

More recent work has explored semi-supervised and hybrid approaches to improve performance with limited labeled data. Zhou *et al.* [5] proposed ECGMatch, a multi-label semi-supervised model for simultaneous prediction of multiple CVDs from ECG recordings with limited supervision, demonstrating robust cross-dataset generalizability. Naz *et al.* [9] proposed a meta-ensemble stacking framework with explainable AI components for heart disease prediction, reporting state-of-the-art performance on benchmark datasets. Orbay and Wikner [4] examined gender-bias attenuation in ML algorithm selection, highlighting that standard model comparison frameworks fail to account for sex-specific predictive utility.

2.2 Data Quality as a Determinant of Healthcare AI Performance

Al-Hgaish *et al.* [2] proposed a robust prioritization framework using the Analytic Hierarchy Process (AHP) and sensitivity analysis to systematically rank healthcare data quality dimensions—accuracy, completeness, consistency, timeliness, and precision—according to their impact on ML-driven healthcare outcomes. Empirical findings demonstrated that accuracy and completeness consistently emerge as the most influential dimensions, with even minor degradation leading to disproportionate declines in model performance regardless of algorithmic sophistication.

2.3 Feature Selection and Model Robustness

Sreehari and Babu [8] conducted a systematic review of feature selection techniques across interdisciplinary domains, demonstrating that filter-based, wrapper-based, and embedded selection methods enhance model efficiency and stability when applied to well-curated datasets. Comparative analyses consistently show that models trained on high-quality datasets outperform those trained on larger but lower-quality alternatives, reinforcing the argument that data quality optimization should precede algorithmic optimization in safety-critical healthcare applications.

2.4 Bias, Fairness, and Equity in Healthcare AI

Raza *et al.* [6] conducted a narrative review exploring bias and prediction metrics for equity-centered public health decision-making, demonstrating that data imbalance, historical underrepresentation, and inappropriate evaluation metrics systematically produce inequitable outcomes in clinical decision support. Gender bias is a persistent issue: models trained on male-dominant datasets frequently underperform for women, resulting in delayed diagnosis and unequal clinical utility [4]. Despite growing awareness, fairness-aware methodologies are rarely integrated as core design principles in cardiovascular AI pipelines.

2.5 Privacy-Preserving and Decentralized Learning

Wei *et al.* [10] proposed DeFedHDP, a fully decentralized online federated learning framework for heart disease prediction that enables collaborative model training without centralized data sharing, addressing privacy, governance, and scalability challenges in distributed healthcare environments. While promising for fragmented LMIC health systems, most federated learning studies prioritize accuracy and communication efficiency over demographic fairness or subgroup-specific performance evaluation.

2.6 Research Gap

Despite substantial progress, persistent gaps remain. Menopausal status is not treated as a distinct cardiovascular risk modifier in existing ML models. Most frameworks rely on high-income datasets, limiting applicability in LMIC contexts. Algorithm-centric evaluations neglect systematic data quality assessment, and fairness-aware methods are rarely core design elements. The present work is motivated by these interconnected gaps.

3 Methodology

3.1 Dataset

The NHANES (National Health and Nutrition Examination Survey) 2017–March 2020 Pre-Pandemic dataset was used as the primary data source. NHANES provides comprehensive health data collected through interviews, physical examinations, and laboratory tests. Key data modules integrated in this study include the demographics module (P_DEMO.XPT), reproductive health questionnaire (P_RHQ.XPT), blood pressure examination (BPX), cholesterol laboratory results (LBXTC, LBDHDD, LBXLDL), and self-reported medical conditions (MCQ). All modules were merged using the unique participant identifier SEQN. Analysis was restricted to female participants to ensure menopausal-stage modeling specificity [4], [6]. Given the absence of publicly available LMIC-specific menopause–CVD datasets, NHANES was employed as a proxy base with menopausal stage simulation and synthetic socioeconomic augmentation designed to reflect LMIC demographic distributions. This simulation-based augmentation strategy is consistent with approaches in cardiovascular ML research that leverage surrogate and augmented datasets to address data scarcity [3], [5].

3.2 Menopause-Specific Feature Engineering

Menopausal status was approximated using clinically accepted age-based thresholds consistent with WHO and ACOG guidelines, as no explicit hormonal measurements were available in the dataset:

- **Stage 0 — Pre-menopause:** Age < 45 years
- **Stage 1 — Peri-menopause:** Age 45–54 years
- **Stage 2 — Post-menopause:** Age ≥ 55 years

Additional menopause–cardiovascular interaction features were engineered to capture hormonal amplification of established risk factors:

meno_bp_risk = menopause_stage × systolic BP ------------(1)

meno_chol_risk = menopause_stage × total cholesterol ------(2)

3.3 LMIC Equity Simulation

To reflect healthcare disparities characteristic of LMIC environments, two socioeconomic proxy variables were introduced following equity-centered public health frameworks [6]:

- **Income Level:** Derived from the family poverty income ratio (INDFMPIR). Participants below 1.85 were classified as low-income (0); above as middle-income (1), approximating a 70/30 LMIC distribution.
- **Healthcare Access:** Conditionally assigned based on income level, with lower-income participants assigned poorer access probabilities (0 = poor, 1 = moderate, 2 = good).

3.4 Data Quality Prioritization

Following the AHP-informed data quality framework of Al-Hgaish *et al.* [2], systematic data quality assessment was performed prior to modeling. Missing values in key clinical variables were handled through median imputation for continuous features and mode imputation for categorical variables. Consistency checks were applied to ensure physiologically plausible ranges for blood pressure, cholesterol, and age-based menopausal classification.

3.5 Machine Learning Model

A Logistic Regression model served as the interpretable baseline. A Random Forest classifier was implemented as the primary predictive model, selected due to its robustness in handling nonlinear interactions, tabular health data, and class-imbalanced distributions [1]. A total of 400 estimator trees were used with class_weight='balanced' to address the severe imbalance between CVD-positive and non-CVD cases.

3.6 Threshold Optimization

The conventional 0.5 classification threshold was replaced with an automatically optimized threshold based on F1-score maximization across the range $[0.2, 0.9]$. The optimal threshold of $\theta = 0.60$ was identified through this automated selection mechanism, improving sensitivity for minority CVD cases while preventing excessive false positives.

3.7 Evaluation Framework

Model performance was evaluated using a comprehensive set of metrics prioritized for imbalanced healthcare data [2], [6]:

- **Accuracy:** overall classification correctness.
- **Recall / Sensitivity:** proportion of true CVD cases detected (*primary metric*).
- **Precision:** proportion of predicted CVD cases that were true positives.
- **F1-score:** harmonic mean of precision and recall.
- **ROC-AUC:** area under the receiver operating characteristic curve, measuring discriminative performance independent of threshold.

4 Experimental Setup

The dataset was partitioned into training and testing sets using an 80/20 stratified split. Stratification was performed on the CVD outcome variable to ensure both sets contained proportional positive and negative cases, preventing data leakage from class-imbalanced distributions. All simulated socioeconomic features and menopause interaction terms were generated prior to splitting to preserve internal consistency.

Three model configurations were evaluated to isolate the contribution of individual modeling choices:

1. **Logistic Regression:** interpretable baseline model with default parameters.
2. **Random Forest (Default):** standard configuration without class-weight adjustment.
3. **Random Forest (Balanced + Threshold-Tuned):** 400 trees, class_weight='balanced', optimal threshold $\theta = 0.60$.

The dataset exhibited substantial class imbalance characteristic of real-world cardiovascular screening: 307 non-CVD cases (97.5%) and 8 CVD-positive cases (2.5%) in the test partition. This extreme skew necessitates sensitivity-focused evaluation metrics, as raw accuracy can be misleading in such distributions (**paul2024?**; **zhou2024semi?**).

5 Results and Analysis

This section presents the experimental results for the three model configurations evaluated on the NHANES 2017–2020 female sub-cohort. Performance is assessed across overall classification metrics, sensitivity and discriminative capability, class-level behavior, menopause-stage subgroup equity, and feature importance. The optimized Random Forest classifier (class-weight balanced, threshold-tuned at $\theta = 0.60$) achieved an accuracy of 89.2%, a recall of 0.875, and a ROC-AUC of 0.914, demonstrating strong predictive performance under severe class imbalance conditions.

5.1 Overall Model Performance Comparison

Table 1

Model Performance Comparison

Model	Accuracy	Recall	ROC-AUC
Logistic Regression	~0.70	~0.40	~0.80
Random Forest (Default)	0.930	0.380	0.914
RF (Balanced + Tuned)	**0.892**	**0.875**	**0.914**

The baseline Logistic Regression model achieved approximately 70% accuracy with a recall of 0.40, establishing an interpretable reference point. The standard Random Forest model demonstrated improved accuracy (93%) but critically low recall (0.38), confirming that majority-class bias dominates default classifiers under severe imbalance. The final balanced, threshold-optimized Random Forest produced the strongest performance profile: 89.2% accuracy, 87.5% recall, and 0.914 ROC-AUC. The modest reduction in accuracy from the default to the balanced configuration reflects the deliberate sensitivity–specificity tradeoff achieved through class-weight balancing and threshold adjustment.

5.2 Sensitivity and ROC-AUC Analysis

The optimized model achieved a recall of 0.875, detecting 7 out of 8 true CVD-positive cases in the test set. In clinical preventive screening, sensitivity is the primary performance criterion: false negatives—missed cardiovascular cases—lead to delayed intervention and increased morbidity risk. This result demonstrates that class-weight balancing and threshold optimization effectively mitigated majority-class bias without sacrificing predictive stability.

The ROC-AUC score of 0.914 confirms excellent discriminative capability, indicating that the classifier assigns substantially higher probability scores to true CVD cases relative to non-CVD cases across varying decision thresholds. An AUC above 0.90 represents strong class separability in binary classification tasks. The optimal classification threshold of 0.60 was identified through automated F1-score maximization, improving minority-class detection while preventing excessive false positives.

5.3 Class-Wise Performance

Table 2

Class-Wise Performance Metrics (RF Balanced + Tuned)

Class	Precision	Recall	F1-Score	Support
Non-CVD (Majority)	1.00	0.89	0.94	307
CVD-Positive (Minority)	0.17	0.88	0.29	8

The majority (non-CVD) class achieved near-perfect precision (1.00) and strong recall (0.89) with an F1-score of 0.94. For the minority (CVD-positive) class, precision is relatively low at 0.17, which is expected given the extreme scarcity of positive samples ($n = 8$). Under such imbalance, even a limited number of false positives substantially reduces precision. However, in the screening context, this tradeoff is clinically acceptable: false positives are addressable through secondary clinical evaluation, whereas false negatives represent missed disease with potentially severe consequences. The recall of 0.88 for the CVD-positive class confirms the system's utility as a first-stage screening tool.

5.4 Menopause-Stage Subgroup Evaluation

Subgroup prediction accuracy was evaluated separately across menopausal stages to assess the framework's equity properties:

- **Stage 0 (Pre-menopausal):** Accuracy = 1.00
- **Stage 1 (Peri-menopausal):** Accuracy = 1.00
- **Stage 2 (Post-menopausal):** Accuracy = 0.824

The performance gradient across menopausal stages aligns with clinical evidence that post-menopausal women carry elevated and more complex cardiovascular risk profiles due to sustained estrogen deprivation and associated metabolic changes. The reduced accuracy in Stage 2 reflects the greater heterogeneity in post-menopausal CVD risk, where multiple interacting physiological pathways—dyslipidemia, hypertension, and insulin resistance—create more varied risk presentations. This finding has direct implications for healthcare resource allocation in LMIC settings, where post-menopausal women represent the highest-risk yet most underserved segment of the cardiovascular screening population.

5.5 Feature Importance Analysis

Table 3

Feature Importance Scores

Feature	Description	Importance
meno_chol_risk	Menopause × Cholesterol	0.169
meno_bp_risk	Menopause × Systolic BP	0.137
RIDAGEYR	Age	0.129
LBXTC	Total Cholesterol	0.110
LBXLDL	LDL Cholesterol	0.097
BPXDI1	Diastolic Blood Pressure	0.071
LBDHDD	HDL Cholesterol	0.069
BPXSY1	Systolic Blood Pressure	0.064
income_level	Income Level (LMIC proxy)	0.059
meno_stage	Menopause Stage	0.052
healthcare_access	Healthcare Access (LMIC proxy)	0.043

The two highest-ranked features are the engineered menopause-specific interaction variables: meno_chol_risk (0.169) and meno_bp_risk (0.137). This result empirically validates the clinical rationale for menopause-aware feature engineering: the amplifying effect of estrogen decline on cholesterol and blood pressure constitutes the strongest learnable cardiovascular risk signal in this cohort. Age (0.129) and total cholesterol (0.110) follow as established cardiovascular risk factors. Notably, the socioeconomic proxy variables—income level (0.059) and healthcare access (0.043)—contribute meaningfully despite their simulated nature, supporting the equity-focused framing of the model.

5.6 Summary of Key Findings

The following findings emerge from the experimental evaluation:

Finding 1 — Class-balancing is essential under severe imbalance. The default Random Forest achieved 93% accuracy but only 0.38 recall, demonstrating that majority-class bias renders standard classifiers ineffective for minority CVD detection. The balanced, threshold-optimized configuration recovered recall to 0.875 while maintaining 89.2% accuracy.

Finding 2 — Menopause-specific interaction features are the strongest CVD predictors. Feature importance analysis ranked meno_chol_risk (0.169) and meno_bp_risk (0.137) as the top two predictors, surpassing established clinical markers including age

(0.129) and total cholesterol (0.110). This empirically validates the clinical rationale for menopause-aware feature engineering.

Finding 3 — Post-menopausal women present the most complex risk profile. Subgroup evaluation revealed perfect classification accuracy (1.00) for pre- and peri-menopausal cohorts, with a significant performance decline for post-menopausal women (0.824), consistent with the greater heterogeneity of sustained estrogen-deficient cardiovascular risk.

Finding 4 — Socioeconomic equity variables contribute meaningfully to prediction. The LMIC-representative proxy variables income_level (0.059) and healthcare_access (0.043) contributed positively to model performance despite their simulated nature, supporting the equity-focused design rationale.

Finding 5 — ROC-AUC of 0.914 confirms strong discriminative capability. The optimized model's ROC-AUC exceeds the 0.90 threshold conventionally indicative of strong class separability in binary clinical classification, confirming robust utility for population-level screening pipelines.

Table 4

Consolidated Results Summary

Metric / Feature	**LR Baseline**	**RF Default**	**RF Balanced + Tuned**
Accuracy	~0.70	0.930	**0.892**
Recall (Sensitivity)	~0.40	0.380	**0.875**
ROC-AUC	~0.80	0.914	**0.914**
Top Feature (Importance)	—	—	**meno_chol_risk (0.169)**
Post-menopausal Subgroup Accuracy	—	—	**0.824**

Taken together, these findings confirm that the proposed menopause-aware, equity-focused framework achieves clinically meaningful cardiovascular risk prediction under conditions representative of LMIC healthcare deployment: extreme class imbalance, absence of hormonal biomarkers, and reliance on standard clinical and demographic variables. The consolidated results are summarized in Table 4.

6 Discussion

The experimental results collectively demonstrate that menopause-aware feature engineering substantially enhances cardiovascular risk prediction for women under realistic healthcare constraints. The two highest-ranked predictive features were both menopause–interaction variables, confirming that the clinical relationship between menopausal hormonal transition and CVD risk is learnable from population health data through appropriately engineered proxies. This finding directly addresses the literature gap identified by Orbay and Wikner [4] regarding the absence of gender-specific life-stage variables in cardiovascular AI.

The data quality prioritization approach, informed by Al-Hgaish *et al.* [2], proved foundational to model performance. By systematically addressing accuracy and completeness—the highest-priority data quality dimensions—prior to algorithmic modeling, the framework avoided the common error of attributing performance gains to algorithmic sophistication when underlying data quality is the true driver [2]. This data-centric perspective distinguishes the present work from algorithm-comparison-focused cardiovascular ML studies.

The integration of socioeconomic indicators reflects the equity-focused objective of the framework. In LMIC settings, income and healthcare access are significant mediators of cardiovascular outcomes, affecting preventive care uptake, diagnostic availability, and treatment adherence [1], [6]. Although the simulated equity variables carry relatively lower feature importance than clinical predictors, their inclusion enables subgroup-level fairness evaluation consistent with the equity-centered public health frameworks recommended by Raza *et al.* [6].

The performance gradient across menopausal stages—with post-menopausal women achieving lower accuracy (0.824) than pre- and peri-menopausal groups (1.00)—highlights the complexity of post-menopausal cardiovascular risk profiling. This observation reinforces the need for specialized, stage-specific prediction components or larger training datasets with confirmed post-menopausal CVD cases. It also carries direct implications for healthcare resource allocation in LMIC systems: post-menopausal women represent the segment of greatest cardiovascular risk, yet standard screening tools—which ignore menopausal stage—are least capable of accurately identifying risk in this population.

The low precision for the minority CVD class (0.17), while suboptimal in isolation, must be interpreted within the clinical screening context. In first-stage population screening, the priority is to maximize recall—ensuring no high-risk cases are missed—while accepting a higher rate of false positives that will be subsequently triaged through secondary clinical evaluation. The precision-recall tradeoff achieved here is clinically appropriate for a preventive screening tool operating at population scale.

The federated and privacy-preserving learning approaches explored in related work [10] represent a promising pathway for scaling the proposed framework to distributed LMIC healthcare networks. Future iterations of the proposed framework should incorporate federated learning mechanisms to enable multi-site training on heterogeneous LMIC datasets without compromising data sovereignty.

7 Limitations

Several limitations of this study are acknowledged. First, menopausal status was approximated using age-based thresholds rather than directly measured hormonal markers (FSH, estradiol) or self-reported reproductive history. This proxy approach, while clinically accepted, introduces classification uncertainty for individuals with atypical menopause timing, including surgical or premature menopause.

Second, the socioeconomic variables representing LMIC healthcare disparities were synthetically simulated rather than directly observed from LMIC data. External validation on real LMIC populations—with confirmed menopausal status, hormonal biomarkers, and socioeconomic indicators—is necessary to establish generalizability.

Third, the extremely low number of CVD-positive female cases in the NHANES 2017–2020 dataset ($n = 8$ in the test set) limits per-stage statistical power and introduces uncertainty in class-level metrics. Larger datasets with confirmed cardiovascular outcomes and menopausal biomarkers would substantially strengthen the findings.

Fourth, feature importance scores derived from Random Forest's mean decrease in impurity may favor continuous variables over binary or categorical features, potentially understating the contribution of socioeconomic indicators. Future work should supplement these with SHAP-based explainability analysis to provide both global and instance-level interpretability.

8 Conclusion

This study presented a sustainable, equity-focused machine learning framework for cardiovascular disease risk prediction among menopausal women, designed for deployment in resource-constrained LMIC healthcare environments. By integrating menopause-stage indicators, menopause–risk interaction features, and socioeconomic proxy variables into a class-balanced Random Forest classifier, the proposed system achieved strong predictive performance under severe class imbalance conditions.

The optimized model achieved 89.2% accuracy, 87.5% recall, and ROC-AUC of 0.914. Feature importance analysis confirmed that menopause–cholesterol and menopause–blood pressure interaction terms were the strongest predictors of cardiovascular risk, empirically validating the clinical rationale for menopause-aware modeling. Subgroup analysis revealed increased predictive complexity for post-menopausal women, consistent with their elevated and heterogeneous cardiovascular risk profile.

These findings confirm that equity-focused AI models incorporating menopause-aware features and LMIC-representative socioeconomic indicators can meaningfully support early cardiovascular risk detection in women, with particular relevance for deployment in settings where preventive screening capacity is limited and postmenopausal women face the greatest unmet burden of cardiovascular disease.

9 Future Work

Future work will focus on validating the proposed framework on real LMIC clinical datasets with confirmed menopause biomarkers and cardiovascular outcomes. Integration of additional equity features—including education level, geographic access, ethnicity, and dietary patterns—will further strengthen the model's fairness properties and LMIC representativeness.

Advanced explainability techniques including SHAP-based global and local interpretation will be incorporated to improve clinical transparency and support regulatory compliance. XGBoost and LightGBM models will be evaluated as alternative classifiers, with particular attention to their performance in extremely imbalanced settings.

Federated learning architectures, following the DeFedHDP framework [10], will be explored to enable privacy-preserving multi-site training across distributed LMIC healthcare networks. Finally, integration of the framework into a mobile health screening application will translate these findings into accessible, deployable preventive healthcare tools for frontline workers in resource-constrained settings.

10 References

[1] H. F. El-Sofany, "Predicting heart diseases using machine learning and different data classification techniques," *IEEE Access*, vol. 12, pp. 106149–106165, 2024, doi: 10.1109/ACCESS.2024.3437181.

[2] A. M. Al-Hgaish, R. B. Atan, R. Yaakob, and M. H. Osman, "A robust prioritization framework of data quality dimensions to improve ML-driven

healthcare systems using AHP and sensitivity analysis," *IEEE Access*, vol. 13, pp. 158059–158070, 2025, doi: 10.1109/ACCESS.2025.3601031.

[3] V. V. Paul and J. A. I. S. Masood, "Exploring predictive methods for cardiovascular disease: A survey of methods and applications," *IEEE Access*, vol. 12, pp. 101497–101503, 2024, doi: 10.1109/ACCESS.2024.3430898.

[4] R. Orbay and E. Wikner, "Sustainable selection of machine learning algorithm for gender-bias attenuated prediction," *IEEE Open J. Veh. Technol.*, 2024, doi: 10.1109/OJVT.2024.3502921.

[5] R. Zhou, L. Lu, Z. Liu, T. Xiang, Z. Liang, D. A. Clifton, Y. Dong, and Y.-T. Zhang, "Semi-supervised learning for multi-label cardiovascular diseases prediction: A multi-dataset study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 5, pp. 3305–3320, May 2024, doi: 10.1109/TPAMI.2023.3342828.

[6] S. Raza, A. Shaban-Nejad, E. Dolatabadi, and H. Mamiya, "Exploring bias and prediction metrics to characterise the fairness of machine learning for equity-centered public health decision-making: A narrative review," *IEEE Access*, vol. 12, pp. 180815–180826, 2024, doi: 10.1109/ACCESS.2024.3509353.

[7] G. Sunilkumar and P. Kumaresan, "Deep learning and transfer learning in cardiology: A review of cardiovascular disease prediction models," *IEEE Access*, vol. 12, pp. 193365–193386, 2024, doi: 10.1109/ACCESS.2024.3514093.

[8] E. Sreehari and L. D. D. Babu, "Feature selection techniques for data analysis and decision making in interdisciplinary areas: A systematic review," *IEEE Access*, vol. 12, pp. 188848–188886, 2024, doi: 10.1109/ACCESS.2024.10781405.

[9] M. Naz *et al.*, "Meta-ensemble learning for heart disease prediction: A stacking-based approach with explainable AI," *IEEE Access*, vol. 13, pp. 137271–137286, 2025, doi: 10.1109/ACCESS.2025.3588683.

[10] M. Wei, J. Yang, Z. Zhao, X. Zhang *et al.*, "DeFedHDP: Fully decentralized online federated learning for heart disease prediction in computational health systems," *IEEE Trans. Comput. Social Syst.*, pp. 1–14, Oct. 2024, doi: 10.1109/TCSS.2024.3406528.

Bibliography:

Hannah Imtiaz is a researcher at the Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, India. Her research interests include machine learning for healthcare, fairness-aware AI, and cardiovascular risk modeling in resource-constrained settings.

Menaha Bhandari is a researcher at the Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, India. Her research interests include health equity, predictive modeling for women's health, and federated learning in distributed healthcare environments.

Dr. Thanigaivelu P. S. received his doctoral degree in Engineering and has served as an Assistant Professor at the Department of Computing Technologies, Faculty of Engineering & Technology, SRM Institute of Science and Technology, Kattankulathur, Chennai, India, for 17 years. He is an experienced educator and researcher with expertise in academic instruction, curriculum development, and student mentoring. His teaching portfolio includes Programming for Problem Solving, Computer Organization and Architecture, Compiler Design, Computer Graphics and Animation, and Advanced Mobile Communications. His research interests span Brain–Computer Interface, Signal Processing, EEG Signals, Machine Learning, Deep Learning, and AI-driven Robotics.