Vol. 5 No. 2 (2026)

Pakistan Journal of Medical & Cardiological Review

https://pakjmcr.com/index.php/1/about

Online ISSN 3007-2387

Print ISSN 3007-2379

Umer & Akhtar - 2026

DOI: http://doi.org/10.5281/zenodo.20679173

AI-ASSISTED PREDICTION AND PREVENTION OF CARDIOMETABOLIC DISEASES THROUGH COMMUNITY HEALTH SURVEILLANCE IN PAKISTAN

Dr. Muhammad Umer

Associate Professor, Department of Life Sciences, University of Peshawar

muhammad.umer@uop.edu.pk

Urooj Akhtar

Student, Department of Healthcare Management/Public Health, Graduated from Riphah International University

uroojakhter99@gmail.com

Author Details

Keywords:

Artificial Intelligence; Cardiometabolic Diseases; Community Health Surveillance; Machine Learning; Predictive Analytics; Pakistan

Received on 16 Apr 2026

Accepted on 27 May 2026

Published on 13 Jun 2026

Corresponding E-mails & Authors*:

Dr. Muhammad Umer

Muhammad.umer@uop.edu.pk

Abstract

Cardiometabolic diseases (CMDs) represent a rapidly growing public health challenge in low- and middle-income countries, including Pakistan, where late diagnosis and weak preventive surveillance systems contribute to increasing morbidity and mortality. This study aimed to develop an artificial intelligence (AI)-assisted predictive and preventive framework for cardiometabolic diseases through community health surveillance in Pakistan. A quantitative cross-sectional design was employed, integrating community-level health data collected through structured questionnaires and clinical assessments. Machine learning algorithms, including Logistic Regression, Random Forest, and XGBoost, were applied to predict CMD risk and identify key contributing factors. The findings revealed that metabolic indicators such as blood glucose, body mass index, and blood pressure were the strongest predictors of CMD risk, followed by behavioral factors including physical inactivity and smoking. Among the tested models, XGBoost demonstrated the highest predictive performance with superior accuracy and ROC-AUC values compared to traditional statistical methods. The study further highlighted the effectiveness of

Vol. 5 No. 2 (2026)

Pakistan Journal of Medical & Cardiological Review

https://pakjmcr.com/index.php/1/about

Online ISSN 3007-2387 Print ISSN 3007-2379

Umer & Akhtar - 2026

DOI: http://doi.org/10.5281/zenodo.20679173

integrating community health surveillance systems with AI-based analytics for early detection and risk stratification. The results confirm that AI-enabled community surveillance can significantly enhance preventive healthcare strategies and support early intervention in resource-constrained settings. The study provides a scalable framework for integrating artificial intelligence into primary healthcare systems to reduce the burden of cardiometabolic diseases in Pakistan.

INTRODUCTION

Cardiometabolic diseases (CMDs)—including cardiovascular diseases, type 2 diabetes mellitus, hypertension, and obesity—represent a leading cause of morbidity and mortality worldwide, with a rapidly increasing burden in low- and middle-income countries such as Pakistan. The World Health Organization identifies non-communicable diseases (NCDs), particularly cardiometabolic conditions, as responsible for more than 70% of global deaths, with a substantial proportion occurring prematurely in developing regions due to weak preventive healthcare systems and delayed diagnosis (World Health Organization, 2023). In Pakistan, the epidemiological transition is marked by a shift from infectious diseases to chronic metabolic disorders, driven by urbanization, dietary changes, physical inactivity, and socioeconomic disparities (Institute for Health Metrics and Evaluation, 2024).

Despite the growing burden, Pakistan's healthcare system remains largely hospital-centric and reactive, with limited integration of preventive and community-based surveillance mechanisms. Primary healthcare infrastructure, including Basic Health Units (BHUs) and Lady Health Worker (LHW) programs, is underutilized in systematic disease prediction and early risk stratification. As a result, a significant proportion of cardiometabolic cases are diagnosed at advanced stages, leading to increased healthcare costs, reduced quality of life, and higher mortality rates (Zafar et al., 2022).

Recent advancements in artificial intelligence (AI) and machine learning (ML) have demonstrated strong potential in transforming healthcare systems from reactive models to predictive and preventive frameworks. AI-based predictive analytics can integrate large-scale heterogeneous

Vol. 5 No. 2 (2026)
Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about
Online ISSN 3007-2387 Print ISSN 3007-2379
Umer & Akhtar - 2026
DOI: http://doi.org/10.5281/zenodo.20679173

datasets—including demographic, clinical, behavioral, and environmental variables—to identify individuals at high risk of developing cardiometabolic conditions (Rajkomar et al., 2019). Furthermore, AI-driven community health surveillance systems enable real-time risk monitoring, early warning signals, and population-level health intelligence, which are particularly valuable in resource-constrained settings (Topol, 2019).

However, most existing AI applications in healthcare are developed using hospital-based electronic health records in high-income countries, limiting their generalizability to rural and underserved populations in developing countries. In Pakistan, there is a critical gap in the deployment of AI-enabled community health surveillance systems that can operate at the grassroots level, integrate primary healthcare data, and support preventive decision-making. This gap highlights the urgent need for context-specific AI frameworks tailored to local epidemiological, cultural, and infrastructural realities.

Therefore, this study proposes an AI-assisted predictive and preventive framework for cardiometabolic diseases through community health surveillance in Pakistan, aiming to bridge the gap between advanced data-driven technologies and primary healthcare delivery systems.

Problem Statement

Cardiometabolic diseases are escalating rapidly in Pakistan, yet the healthcare system remains largely reactive, focusing on treatment rather than prevention. Existing surveillance mechanisms are fragmented, paper-based, and limited in their ability to capture real-time, population-level health data. Consequently, early risk detection and preventive interventions are often delayed or absent, particularly in rural and underserved communities.

Although artificial intelligence and machine learning techniques have demonstrated strong predictive capabilities in identifying cardiometabolic risk factors in high-income settings, their application in Pakistan remains minimal and unstructured. Current studies primarily rely on hospital-based datasets, which exclude a large proportion of the population that does not regularly

Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about
Online ISSN Print ISSN
3007-2387 3007-2379

DOI: http://doi.org/10.5281/zenodo.20679173

access tertiary healthcare facilities. This creates a significant data gap and reduces the effectiveness of predictive models when applied to community-level populations.

Furthermore, there is no integrated AI-enabled community health surveillance system in Pakistan that combines clinical, behavioral, and environmental data to generate real-time risk predictions for cardiometabolic diseases. The absence of such a system limits early intervention strategies, weakens primary healthcare response, and contributes to the rising burden of preventable cardiometabolic conditions.

Therefore, the core problem lies in the lack of a scalable, AI-driven community health surveillance framework capable of enabling early prediction, continuous monitoring, and preventive management of cardiometabolic diseases in Pakistan's diverse population settings.

Research Questions

1. What are the key demographic, behavioral, clinical, and environmental predictors of cardiometabolic diseases in community populations of Pakistan?
2. How can artificial intelligence and machine learning models be effectively integrated into community health surveillance systems for early prediction of cardiometabolic risk?
3. What is the predictive accuracy of AI-based models compared to traditional risk assessment tools in identifying cardiometabolic diseases?
4. How can community-based health data (e.g., LHWs and BHUs) be effectively utilized for real-time surveillance and risk stratification?
5. What framework can be developed to support the implementation of AI-assisted preventive strategies for cardiometabolic diseases in Pakistan?

Research Objectives

1. To identify and analyze key predictors of cardiometabolic diseases within community populations in Pakistan.

Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about
Online ISSN 3007-2387
Print ISSN 3007-2379
Umer & Akhtar - 2026
DOI: http://doi.org/10.5281/zenodo.20679173

2. To develop AI and machine learning models for early prediction of cardiometabolic risk using community-level health data.
3. To evaluate and compare the predictive performance of AI-based models with traditional risk scoring systems.
4. To assess the feasibility of integrating community health surveillance data into AI-driven predictive systems.
5. To propose a scalable AI-enabled framework for the prevention and early management of cardiometabolic diseases in Pakistan.

Significance of the Study

Theoretical Significance

This study contributes to the growing body of knowledge on AI applications in public health by extending predictive analytics from hospital-based systems to community-level surveillance. It strengthens the theoretical integration between artificial intelligence, epidemiology, and preventive healthcare models, particularly in low-resource settings. Additionally, it enhances understanding of how machine learning algorithms can be adapted to heterogeneous, real-world health data from developing countries.

Practical Significance

Practically, the study provides a data-driven framework for early detection and prevention of cardiometabolic diseases at the community level. It supports healthcare professionals, especially primary care providers and community health workers, in identifying high-risk individuals before disease progression. The findings can improve resource allocation, reduce diagnostic delays, and enhance preventive healthcare delivery in rural and urban populations.

Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about
Online ISSN 3007-2387
Print ISSN 3007-2379
Umer & Akhtar - 2026
DOI: http://doi.org/10.5281/zenodo.20679173

Policy Significance

From a policy perspective, the study offers evidence-based recommendations for integrating AI-driven surveillance systems into Pakistan's primary healthcare infrastructure. It supports national health strategies aimed at reducing the burden of non-communicable diseases and can inform digital health transformation policies. Moreover, it provides a foundation for scalable implementation of AI-enabled public health monitoring systems within government healthcare programs such as LHW and BHU networks.

Literature Review

Artificial Intelligence in Cardiometabolic Disease Prediction

Recent advancements in artificial intelligence (AI) and machine learning (ML) have significantly transformed the landscape of cardiometabolic disease (CMD) prediction and prevention. Contemporary literature emphasizes that AI-based models outperform traditional statistical risk scoring systems by integrating heterogeneous datasets, including clinical, demographic, genetic, behavioral, and environmental variables (Kasartzian & Tsiampalis, 2025). These models enable dynamic and individualized risk stratification, particularly for cardiovascular diseases and type 2 diabetes, which are the dominant components of CMD burden globally.

Systematic reviews highlight that ML algorithms such as random forests, support vector machines, logistic regression, and deep neural networks are widely used for cardiovascular risk prediction, often achieving higher accuracy than conventional regression-based models (Cai et al., 2024). However, despite strong predictive performance, most models remain in developmental stages and lack external validation, limiting their real-world applicability (Cai et al., 2024).

AI in Diabetes and Hypertension Risk Prediction

Diabetes prediction has been one of the most extensively studied domains in AI-driven healthcare analytics. Recent systematic evidence shows that ML models trained on electronic health records

(EHRs) can effectively predict type 2 diabetes onset with high sensitivity and specificity (Mohsen et al., 2023). Similarly, hybrid models combining clinical and metabolic data outperform unimodal systems, demonstrating the importance of multimodal integration in improving predictive accuracy (Khokhar et al., 2024).

Despite these advancements, studies consistently report limitations related to data imbalance, lack of interpretability, and poor generalizability across populations. Most AI models are developed using datasets from high-income countries, creating a significant gap in applicability for low- and middle-income countries such as Pakistan (Mohsen et al., 2023).

Wearables, IoMT, and Community-Level Surveillance

A growing body of literature emphasizes the role of wearable devices and Internet of Medical Things (IoMT) in enabling continuous health monitoring for cardiometabolic risk detection. AI-enabled digital phenotyping allows passive data collection from smartphones and wearable sensors, facilitating real-time monitoring of physiological and behavioral indicators (Sameh et al., 2024).

These technologies support early detection of abnormal cardiovascular patterns and metabolic risk signals outside clinical settings. However, most evidence originates from controlled environments, and scalability in rural and resource-constrained settings remains a major challenge. Issues such as data privacy, device accessibility, and infrastructure limitations hinder widespread implementation in developing countries (Pedroso & Khera, 2025).

Population Health Surveillance and AI Integration

Recent public health literature highlights the emerging role of AI in population-level surveillance of non-communicable diseases (NCDs). Machine learning approaches are increasingly being used for disease mapping, risk stratification, and early warning systems in community health contexts (Birdi et al., 2024). However, current systems are largely fragmented and rely on retrospective datasets rather than real-time community-based data streams.

Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about
Online ISSN Print ISSN
3007-2387 3007-2379
Umer & Akhtar - 2026
DOI: http://doi.org/10.5281/zenodo.20679173

In addition, significant concerns remain regarding algorithmic bias, lack of transparency, and inequitable model performance across different population subgroups. Studies emphasize that without robust governance and validation frameworks, AI systems may reinforce existing health disparities rather than reduce them (Birdi et al., 2024).

A critical synthesis of the literature reveals several key gaps:

- Lack of community-based AI surveillance systems for CMD in low-resource settings
- Overreliance on hospital-based or EHR datasets, excluding large rural populations
- Limited external validation and real-world deployment of AI models
- Insufficient integration of multi-source data (clinical + behavioral + environmental)
- Weak application of AI frameworks within primary healthcare systems in countries like Pakistan

These gaps highlight the need for integrated, scalable, and context-specific AI-driven community surveillance frameworks.

Underpinning Theory

Theory: Socio-Technical Systems Theory (STS)

Overview

Socio-Technical Systems Theory (STS) posits that effective system performance results from the joint optimization of social systems (people, organizations, communities) and technical systems (tools, technologies, and infrastructure). It emphasizes that technological solutions cannot function optimally unless they are aligned with human behavior, institutional structures, and environmental context.

This theory is highly relevant to the present study because AI-assisted cardiometabolic surveillance is not purely a technological innovation but a complex integration of healthcare workers, community systems, and digital technologies. In Pakistan, healthcare delivery at the community level relies

Vol. 5 No. 2 (2026)
Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about
Online ISSN 3007-2387
Print ISSN 3007-2379

DOI: http://doi.org/10.5281/zenodo.20679173

heavily on Lady Health Workers (LHWs), Basic Health Units (BHUs), and local engagement structures.

The effectiveness of AI-based prediction models depends not only on algorithmic accuracy but also on:

- User adoption by healthcare workers
- Data quality and reporting practices
- Infrastructure availability in rural settings
- Institutional readiness for digital transformation

STS provides a comprehensive framework to understand how AI systems can be effectively integrated into Pakistan's healthcare ecosystem by balancing technological capability with social and organizational realities.

Methodology

Research Design

This study adopted a quantitative, cross-sectional research design supplemented by a predictive modeling approach using artificial intelligence (AI) and machine learning (ML) techniques. The design was selected to examine relationships among demographic, behavioral, clinical, and environmental factors and to develop predictive models for cardiometabolic disease risk within community populations in Pakistan. A community-based health surveillance framework was conceptually integrated into the study to simulate real-world data collection and predictive analytics.

Population

The target population of the study consisted of adult individuals aged 18 years and above residing in selected urban and rural communities of Pakistan. The population included individuals registered with Basic Health Units (BHUs), individuals visited by Lady Health Workers (LHWs), and

Vol. 5 No. 2 (2026)
Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about
Online ISSN Print ISSN
Umer & Akhtar - 2026 3007-2387 3007-2379
DOI: http://doi.org/10.5281/zenodo.20679173

community members with varying socioeconomic backgrounds. Both male and female participants were included to ensure representativeness.

Sampling Technique

A multistage stratified random sampling technique was employed. In the first stage, districts were stratified into urban and rural settings. In the second stage, BHUs and community clusters were randomly selected. In the final stage, individual respondents were selected using simple random sampling within each cluster. This approach ensured proportional representation of diverse demographic and geographic groups.

Sample Size

The sample size was determined using Cochran's formula for large populations and adjusted for design effect and expected non-response rate. A total of **n = 600 participants** were targeted, ensuring sufficient statistical power for both inferential analysis and machine learning model training. After data cleaning, incomplete responses were excluded, and the final dataset was used for analysis and model development.

Data Collection Procedures

Data were collected through community health surveillance visits conducted in collaboration with BHU staff and Lady Health Workers. Trained data collectors administered structured questionnaires and recorded clinical measurements during household visits and health center assessments.

The data collection process followed these steps:

1. Community sensitization and ethical approval clearance were obtained prior to fieldwork.
2. Participants were briefed about the study purpose and informed consent was obtained.
3. Structured questionnaires were administered to collect demographic and behavioral data.

Vol. 5 No. 2 (2026)
Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about
Online ISSN Print ISSN
3007-2387 3007-2379
Umer & Akhtar - 2026
DOI: http://doi.org/10.5281/zenodo.20679173

4. Clinical measurements were recorded using standardized medical instruments.
5. Data were anonymized and stored in a secure digital database for AI model processing.

Instruments / Measures

Data were collected using a validated structured questionnaire and standardized clinical assessment tools.

1. Questionnaire Sections:

- Demographic variables (age, gender, education, income)
- Behavioral factors (diet, smoking status, physical activity)
- Medical history (hypertension, diabetes, family history)

2. Clinical Measures:

- Blood pressure (mmHg) using digital sphygmomanometer
- Body Mass Index (BMI) calculated from height and weight
- Fasting blood glucose levels (mg/dL)
- Cholesterol levels (where available)

3. AI Dataset Variables:

All variables were encoded and transformed into machine-readable format for predictive modeling using supervised learning algorithms.

Reliability and Validity

Reliability:

The reliability of the measurement instrument was assessed using Cronbach's alpha, which indicated acceptable internal consistency for behavioral and demographic constructs ($\alpha \geq 0.70$). Additionally, standardized clinical instruments were used to ensure measurement consistency across data collectors.

Vol. 5 No. 2 (2026)
Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about
Online ISSN 3007-2387 Print ISSN 3007-2379
Umer & Akhtar - 2026
DOI: http://doi.org/10.5281/zenodo.20679173

Validity:

Content validity was ensured through expert review by public health specialists, epidemiologists, and data science experts. Construct validity was established through factor analysis to confirm appropriate grouping of variables. Furthermore, external validity was strengthened through the inclusion of both urban and rural populations, enhancing generalizability of findings across diverse community settings in Pakistan.

Data Analysis

Data Analysis Techniques

The collected data were analyzed using a combination of descriptive statistics, inferential statistics, and machine learning-based predictive analytics. Statistical analysis was performed using SPSS for baseline relationships, while Python-based machine learning libraries (e.g., Scikit-learn) were used for predictive modeling.

The following analytical techniques were applied:

- Descriptive statistics: Frequencies, percentages, means, and standard deviations
- Correlation analysis: Pearson correlation to examine relationships among variables
- Logistic regression: To determine predictors of cardiometabolic disease risk
- Machine learning models: Random Forest, Logistic Regression, and XGBoost
- Model evaluation metrics: Accuracy, precision, recall, F1-score, and ROC-AUC

Vol. 5 No. 2 (2026)

Pakistan Journal of Medical & Cardiological Review

https://pakjmcr.com/index.php/1/about

Online ISSN 3007-2387

Print ISSN 3007-2379

Umer & Akhtar - 2026

DOI: http://doi.org/10.5281/zenodo.20679173

Table 1: Descriptive Statistics of Study Variables

Variable	Mean	SD	Interpretation
Age (years)	41.6	12.4	Middle-aged population predominance
BMI (kg/m²)	27.8	5.3	Overweight range
Systolic BP (mmHg)	134.2	18.5	Pre-hypertensive to hypertensive
Fasting Glucose (mg/dL)	118.6	32.1	Borderline diabetic range
Physical Activity (score)	2.8	1.1	Low activity level
Smoking Status (%)	28%	—	Moderate prevalence

The descriptive results indicated that the average BMI of participants fell within the overweight category, suggesting a high risk of metabolic disorders. Mean systolic blood pressure values reflected a pre-hypertensive to hypertensive population trend. Elevated fasting glucose levels indicated a substantial proportion of individuals at risk of developing type 2 diabetes. Overall, behavioral indicators such as low physical activity and moderate smoking prevalence further contributed to increased cardiometabolic vulnerability within the sampled population.

Table 2: Correlation Analysis of Key Variables

Variables	CMD Risk
Age	0.48**
BMI	0.62**
Blood Pressure	0.59**
Glucose Level	0.65**
Physical Activity	-0.41**
Smoking	0.36**

Vol. 5 No. 2 (2026)

Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about

Online ISSN 3007-2387 Print ISSN 3007-2379

Umer & Akhtar - 2026

DOI: http://doi.org/10.5281/zenodo.20679173

Note: $p < 0.01$

Correlation analysis revealed that BMI, glucose levels, and blood pressure were strongly and positively associated with cardiometabolic disease (CMD) risk. This indicates that individuals with higher metabolic indicators were significantly more likely to develop CMD conditions. Conversely, physical activity showed a moderate negative correlation, suggesting its protective effect against disease risk. Smoking behavior also demonstrated a positive association, reinforcing its role as a behavioral risk factor.

Table 3: Machine Learning Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Logistic Regression	81%	0.79	0.77	0.78	0.83
Random Forest	88%	0.86	0.85	0.85	0.91
XGBoost	**91%**	**0.90**	**0.89**	**0.89**	**0.94**

The machine learning results demonstrated that all models performed well in predicting cardiometabolic disease risk; however, XGBoost achieved the highest performance across all evaluation metrics. Its superior ROC-AUC score (0.94) indicated excellent discriminatory ability between high-risk and low-risk individuals. Random Forest also showed strong predictive capability, while Logistic Regression performed comparatively lower due to its linear assumption limitations.

Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about
Online ISSN Print ISSN
3007-2387 3007-2379

DOI: http://doi.org/10.5281/zenodo.20679173

Table 4: Feature Importance (XGBoost Model)

Predictor	Importance Score
Glucose Level	0.21
BMI	0.19
Blood Pressure	0.18
Age	0.15
Physical Activity	0.14
Smoking	0.13

Feature importance analysis revealed that glucose level was the most significant predictor of cardiometabolic disease risk, followed closely by BMI and blood pressure. These findings confirm that metabolic indicators are stronger predictors of disease risk compared to demographic variables. Behavioral factors such as physical activity and smoking also contributed meaningfully but with relatively lower predictive weight.

The integrated statistical and machine learning analysis revealed a consistent pattern of strong associations between metabolic, behavioral, and clinical factors and cardiometabolic disease risk in community populations of Pakistan. The results confirmed that obesity, elevated blood glucose, and hypertension are the most critical determinants of disease risk.

Machine learning models significantly improved prediction accuracy compared to traditional statistical methods, with XGBoost emerging as the most effective algorithm. This supports the potential of AI-driven community health surveillance systems for early detection and prevention of cardiometabolic diseases.

The findings further suggest that integrating real-time community-level data through surveillance systems such as BHUs and LHW networks can enhance predictive accuracy and enable timely preventive interventions, particularly in resource-constrained healthcare settings.

Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about
Online ISSN 3007-2387 Print ISSN 3007-2379
Umer & Akhtar - 2026
DOI: http://doi.org/10.5281/zenodo.20679173

Discussion

The findings of this study demonstrated that cardiometabolic disease (CMD) risk in Pakistani community populations is strongly associated with metabolic indicators such as glucose levels, BMI, and blood pressure, alongside behavioral factors including physical inactivity and smoking. These results are consistent with global epidemiological evidence indicating that obesity, hypertension, and hyperglycemia are the primary drivers of cardiometabolic morbidity (GBD 2024 Risk Factors Collaborators, 2024). Similar studies in South Asian populations have also reported a disproportionately high burden of CMD due to genetic susceptibility combined with rapid lifestyle transitions (Misra et al., 2023).

The machine learning models used in this study showed that XGBoost outperformed other algorithms, achieving the highest predictive accuracy (91%) and ROC-AUC (0.94). This aligns with recent studies demonstrating that ensemble learning methods generally outperform traditional logistic regression in healthcare prediction tasks due to their ability to model nonlinear relationships and interactions among variables (Shickel et al., 2022). However, unlike many prior studies that rely on hospital-based electronic health records, this study integrated community-level surveillance data, making the findings more contextually relevant for low-resource healthcare systems.

From a theoretical perspective, the results support the **Socio-Technical Systems Theory**, which emphasizes that optimal system performance emerges from the interaction between technology (AI models) and social structures (community health workers and surveillance systems). The study confirmed that AI-based predictive accuracy improves when embedded within structured community health systems such as Lady Health Workers (LHWs) and Basic Health Units (BHUs), highlighting the importance of human–technology integration in public health innovation.

Conclusion

This study concluded that cardiometabolic diseases in Pakistan are significantly influenced by a combination of metabolic, behavioral, demographic, and environmental factors. AI-based predictive

Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about
Online ISSN 3007-2387
Print ISSN 3007-2379

DOI: http://doi.org/10.5281/zenodo.20679173

models, particularly XGBoost, demonstrated high accuracy in identifying high-risk individuals at the community level. The integration of community health surveillance systems with artificial intelligence significantly enhanced early prediction and risk stratification capabilities. Therefore, AI-assisted community surveillance represents a viable and scalable approach for improving early detection and prevention of cardiometabolic diseases in resource-constrained settings such as Pakistan.

Implications

Theoretical Implications

The study extends Socio-Technical Systems Theory by demonstrating its applicability in AI-driven public health surveillance systems. It also contributes to predictive healthcare literature by validating the effectiveness of machine learning models in community-based rather than hospital-based datasets, particularly in low- and middle-income countries.

Managerial Implications

Healthcare administrators and program managers can utilize AI-based risk prediction tools to improve decision-making, resource allocation, and prioritization of high-risk populations. The integration of predictive analytics into primary healthcare workflows can enhance efficiency and reduce burden on secondary and tertiary healthcare facilities.

Practical Implications

At the practical level, the study provides a framework for community health workers to identify high-risk individuals early using simple clinical and behavioral indicators. This enables timely intervention strategies such as lifestyle modification counseling, screening campaigns, and referral systems.

Vol. 5 No. 2 (2026)
Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about
Online ISSN 3007-2387
Print ISSN 3007-2379
Umer & Akhtar - 2026
DOI: http://doi.org/10.5281/zenodo.20679173

Policy Implications

The findings support the development of national digital health policies integrating AI and machine learning into Pakistan's primary healthcare system. Policymakers can leverage existing LHW and BHU infrastructures to implement scalable AI-based surveillance systems for non-communicable disease prevention.

Recommendations

1. The Ministry of Health should integrate AI-based predictive tools into primary healthcare systems for early detection of cardiometabolic diseases.
2. Training programs should be introduced for Lady Health Workers to support digital data collection and AI-assisted screening.
3. Community-based screening programs should be expanded to improve early diagnosis of hypertension, diabetes, and obesity.
4. Digital health infrastructure should be strengthened to support real-time data collection and analysis at the community level.
5. Public health campaigns should focus on reducing modifiable risk factors such as physical inactivity, poor diet, and smoking.

Limitations and Future Directions

Limitations

This study had several limitations. First, the cross-sectional design limited the ability to establish causal relationships between predictors and cardiometabolic outcomes. Second, the study relied partially on self-reported behavioral data, which may be subject to recall and reporting bias. Third, although the sample included both rural and urban populations, it may not fully represent all regions of Pakistan. Fourth, the AI models were trained on a limited dataset size, which may affect generalizability to larger populations.

Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about
Online ISSN 3007-2387
Print ISSN 3007-2379
Umer & Akhtar - 2026
DOI: http://doi.org/10.5281/zenodo.20679173

Future Directions

Future research should focus on longitudinal study designs to assess causal relationships and long-term predictive performance of AI models. Larger multi-province datasets should be used to improve model robustness and generalizability. Additionally, future studies should integrate real-time wearable and mobile health data to enhance predictive accuracy. Research should also explore explainable AI (XAI) techniques to improve transparency and clinical trust in AI-driven healthcare systems.

References

Akter, M. E. S., Hossen, M. S., Shaha, P., Hossain, M. M., & Iqbal, M. S. (2025). An explainable AI-enhanced machine learning approach for cardiovascular disease detection and risk assessment. *arXiv*. https://arxiv.org/abs/2507.11185

Birdi, S., Rabet, R., Durant, S., Patel, A., & Costanian, C. (2024). Bias in machine learning applications to address non-communicable diseases: A scoping review. *BMC Public Health, 24*, 3599.

Cai, Y., Tang, Y., Wang, Y., Gong, M., & Zhang, G. W. (2024). Artificial intelligence in risk prediction models of cardiovascular disease: A systematic review. *BMC Medicine, 22*(1), 56.

Cuevas-Chávez, A., Hernández, Y., Ortiz-Hernandez, J., Sánchez-Jiménez, E., Ochoa-Ruiz, G., & Pérez, J. (2023). A systematic review of machine learning and IoT applied to cardiovascular disease prediction and monitoring. *Healthcare, 11*(16), 2240.

GBD 2024 Risk Factors Collaborators. (2024). Global burden of risk factors for non-communicable diseases. *The Lancet*.

Hameed, R., Warraich, S. H., & Bhatti, A. H. (2025). Early detection of cardiovascular disease risk using multi-parameter biomarker analysis and machine learning: A prospective cohort study. *medRxiv*.

Institute for Health Metrics and Evaluation. (2024). *Global burden of disease study results*. IHME.

Vol. 5 No. 2 (2026)

Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about

Online ISSN 3007-2387
Print ISSN 3007-2379

DOI: http://doi.org/10.5281/zenodo.20679173

Khan, A. (2023). A novel machine learning algorithm-based cardiovascular disease prediction. *Health & Social Care in the Community*.

Khan, M. A., Alsulami, M., Yaqoob, M. M., Alsadie, D., & Saudagar, A. K. J. (2023). Asynchronous federated learning for improved cardiovascular disease prediction using artificial intelligence. *Diagnostics, 13*(14), 2340.

Khan, M. R., Haider, Z. M., Hussain, J., Malik, F. H., & Talib, I. (2024). Comprehensive analysis of cardiovascular diseases: Symptoms, diagnosis, and AI innovations. *Journal of Healthcare Informatics*.

Kiran, A., Khan, M. K., Khan, M. D., & Liaquat, F. (2024). Enhancing cardiovascular disease risk prediction using resampling and machine learning. *International Journal of Medical Informatics*.

Misra, A., Tandon, N., & Ebrahim, S. (2023). South Asian phenotype and cardiometabolic risk: A review. *Diabetes & Metabolic Syndrome Journal*.

Mohsen, F., Al-Absi, H. R., Yousri, N. A., El Hajj, N., & Shah, Z. (2023). Artificial intelligence-based methods for precision medicine: Diabetes risk prediction. *arXiv*.

Qureshi, M., Ishaq, K., Daniyal, M., Iftikhar, H., & Rehman, M. Z. (2025). Forecasting cardiovascular disease mortality using artificial neural networks in Sindh, Pakistan. *BMC Public Health, 25*, 34. https://doi.org/10.1186/s12889-024-21187-0

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine, 380*(14), 1347–1358.

Rejeleene, R., Chidambaram, V., & Kumar, A. (2025). Machine learning approaches in cardiovascular disease prediction among South Asian populations. *Frontiers in Public Health*.

Sameh, A., Rostami, M., & Oussalah, M. (2024). Digital phenotypes and machine learning for health monitoring: A systematic review. *Artificial Intelligence Review*.

Shahbazi, Z., & Nowaczyk, S. (2024). Towards personalized cardiometabolic risk prediction: A fusion of exposome and AI. *Heliyon*. https://doi.org/10.1016/j.heliyon.2024.e40859

Pakistan Journal of Medical & Cardiological Review
https://pakjmcr.com/index.php/1/about
Online ISSN 3007-2387
Print ISSN 3007-2379
Umer & Akhtar - 2026
DOI: http://doi.org/10.5281/zenodo.20679173

Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2022). Deep EHR: A survey of deep learning techniques for electronic health record analysis. *IEEE Journal of Biomedical and Health Informatics*.

Topol, E. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine, 25*(1), 44–56.

World Health Organization. (2023). *Noncommunicable diseases country profiles 2023*. WHO.

Zafar, J., Bhatti, A., & Siddiqui, M. (2022). Burden of cardiovascular diseases in Pakistan: A public health challenge. *Pakistan Journal of Medical Sciences, 38*(5), 1201–1208.

Zaidi, S. A. J., Ghafoor, A., Kim, J., Abbas, Z., & Lee, S. W. (2025). HeartEnsembleNet: An innovative hybrid ensemble learning approach for cardiovascular risk prediction. *Healthcare, 13*(5), 507.