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

Early Detection of CVDs (Biomarkers)

Ms. Suhani Methi

Student

Abstract
This piece discusses the significance of early detection of cardiovascular diseases (CVDs) through novel biomarkers identified using multi-omics data integrated with artificial intelligence (AI) and machine learning (ML). It highlights the challenges in diagnosis, the potential of blood-borne biomarkers, and innovative methods like microfluidics and advanced imaging to enhance clinical practice and patient outcomes.

OVERVIEW:
Cardiovascular diseases are one of the leading causes of mortality worldwide. One study had been conducted, its objective was aimed to identify novel biomarkers and improve disease prediction for cardiovascular diseases (CVDs) using multi-omics data (transcriptomics, genomics, and clinical demographics) combined with artificial intelligence (AI) and machine learning (ML)

Detecting CVDs
The identification and prognosis of the potential for developing Cardiovascular Diseases (CVD) in healthy individuals is a vital aspect of disease management. Accessing the comprehensive health data on CVD currently available within hospital databases holds significant potential for the early detection and diagnosis of CVD, thereby positively impacting disease outcomes.disease The incorporation of machine learning methods holds significant promise in the advancement of clinical practice for the management of Cardiovascular Diseases (CVDs).
A blood borne biomarker is an alternative to imaging or stress testing as it would be cheaper and logistically faster to obtain. A number of candidate biomarkers have been proposed for the detection of cardiac ischemia, however only ischemia modified albumin has been released for clinical use. IMA is a good discriminator between ischemic and non-ischemic patients. Changes in IMA concentration have shown to occur during coronary angioplasty-induced ischemia. Clinical studies indicate that IMA appears to offer on admission an early test which can be combined with electrocardiographic findings and cardiac troponin measurements for the early exclusion of acute coronary syndrome. IMA is an independent predictor of short and long term adverse outcomes in patients with acute chest pain. However, this test is relatively new and uncertainties remain. Elevations of IMA occur in conditions other than chest pain, thus questioning its specificity. The mechanism of IMA formation and the precise entity being measured are not fully known. Nevertheless, IMA measurement remains the only current clinical biomarker which may be used for the diagnosis of patients suspected of cardiac ischemia
Biomarkers for Ischemic Heart Disease include lipids, cholesterol, calcium, and cellular debris Biomarkers for myocardial infarction include severe blood pressure changes, central chest pain.
Recent advantages in highly multiplexed immunoassays have allowed systematic large scale measurement of hundreds.

Additional biomarkers in early CVD detection:

- High-Sensitivity Troponin (hs-TnT/I): Detects myocardial injury at lower thresholds, aiding in early diagnosis of acute coronary syndromes (ACS)
- B-type Natriuretic Peptide (BNP/NT-proBNP): Predicts heart failure risk before symptomatic onset
- C-reactive protein (CRP): Indicates systemic inflammation linked to atherosclerosis
- Lipoprotein(a) [Lp(a)]: An emerging genetic marker for premature coronary artery disease

A blood borne biomarker is an alternative to imaging or stress testing as it would be cheaper and logistically faster to obtain. A number of candidate biomarkers have been proposed for the detection of cardiac ischemia, however only ischemia modified albumin has been released for clinical use. IMA is a good discriminator between ischemic and non-ischemic patients. Changes in IMA concentration have shown to occur during coronary.

Microfluidics is an important part of the process of early detection which has been majorly seen being used in the upcoming field of biomedical engineering. It is first made through being placed in a mold with a silicon base mixed thoroughly with a curating liquid to form a solvent. This is then placed in a vacuum in order to remove all air bubbles and ensure that the solution is free of them in order for it to be placed into the mold to create it. After that, a microfluidics pump is used to push a syringe at a certain flow rate, volume, or diameter in order to control how much of the liquid is going through the microfluidics chip. The chip can then be used to separate the cells from the sample and also the DNA, which can be tested and checked for biomarkers that can be evident of certain diseases. These separate DNA which are filtered out of the sample are key to the upcoming field of BME as early detection of diseases can become feasible and much more efficient. Magnetic chips can also be placed inside these chips where magnetic beads are inserted into the sample which have a positive charge, while the contents of the sample are filtered to separate, as said before, the DNA from the living sample.

In the process of making the microfluidic chips, this includes developing a CAD model of different plausible designs of the chip in order to have a variety to test to see which fits the best. These CAD models are sent to a 3D printer which produces the measurements listed in the CAD model accurately and to the desire of the maker. The use of the printer then ensures that the chip mold is made completely accurate and that it is completely functional and can be used for the processes needed and produce accurate results.

One study conducted:

Methodology used:

1. **RNA sequencing and whole genome sequencing data** collected from 71 patients, 61 CVD patients and 10 controls
2. **(DESeq2) Differential Expression analysis for sequence count data** is a statistical process used to identify genes which are upregulated (over expressed) or down regulated (under expressed)
   Widely trusted process for RNA sequence analysis
   In this specific study, DESeq2 identified 28 differentially expressed genes (DEGs)

Minimum redundancy and maximum relevance (MRMR) used to select 10 relevant genes and 17 single nucleotide polymorphisms which when present show a 100% likelihood to develop cardiovascular diseases (MRMR is a feature selection method used in bioinformatics and machine learning to select the most useful variables and eliminate redundancy so that research is efficient, to understand this in another context, can be used in cancer diagnosis to identify a subset of genes which are most relevant to the onset of cancer but minimise correlations with each other)

3. Pathogenic virus analysis

17 harmful single nucleotide polymorphisms were detected using **Combined Annotation Dependent Depletion (CADD)** and **gnomAD**

CADD is a bioinformatics tool that predicts how harmful a genetic variant (eg. single nucleotide polymorphisms) by integrating diverse genomic annotations (conservation, functional impact, etc.) into a single score.

CADD outputs a PHRED score, from 1-20, the higher the more pathogenic the mutation is

If the score is >20, it is in the top 1% of dangerous mutations.

This measuring technique is used to identify harmful mutations, used in this case to detect the 17 harmful SNPs

gnomAD is a public database of over 140,000 logged genetic variants, these are used to classify whether a mutation is rare.

This **gnomAD** technique is combined with that of **CADD** in order to pinpoint only the most rare and most harmful mutations which can lead to CVDs from all mutations that were recorded

4. Multimodal Artificial Intelligence/Machine Learning Analysis

Trained artificial intelligence in order to be able to predict diseases

Models such as RandomForest, XGBoost, Logistic Regression (LR) trained using multi-omics data

XGBoost found 100% clinical success in identifying CVD Disease in comparison to controls

Key findings of study:

- 7/10 selected genetic biomarkers were previously linked to CVDs
- 17 SNPs were detected as pathogenic with high CADD scores found in regulatory regions of CVD-associated genes

Multiple clinical applications of this research

Estimated number of cardiovascular disease patients in China alone are **290 million** in 2017 with deaths from CVDs accounting for almost **40%** of all deaths, important to take action on this

SCIENTIFIC CHALLENGES FACED

Early identification of cardiovascular diseases is important to improve patient results, however, there are considerable scientific obstacles in making the diagnosis timely and accurate. To begin with, there is a lack of awareness and understanding of the condition, and negative associations with the clinical identifiers. To further elaborate, the key challenge is the lack of highly sensitive and specific biomarkers. Existing biomarkers such as troponins and B-type natriuretic peptide (BNP) are useful when evaluating patients with advanced disease, but are not sensitive enough to detect CVD at the early stages of the disease. Additionally, biomarkers can be challenged by false positives and false negatives related to interference from comorbidities, such as renal disease. While newer biomarkers like microRNAs and inflammatory mediators may have some relevance, there is still a considerable distance to treat them regularly in practice because it may take some time for them to be fully validated in population studies.

Another significant barrier is the limitation of subclinical atherosclerosis. CAD imaging methods provide useful information, but there are constraints. First, imaging requires expensive technology and specialized personnel, which is not available in many clinical situations. Second, imaging primarily identifies calcified plaques, and misses non-calcified high-risk lesions that lead to acute coronary events, a significant gap that calls for more sophisticated imaging technology to identify which plaques would likely rupture. The heterogeneity of CVD also limits the potential to find CVD in its earliest stages. CVD comprises a wide range of disease states, including coronary artery disease, heart failure, and arrhythmias, with each having its own pathophysiologic mechanism. Serious limitations exist with all forms of contemporary screening tools, which cannot tell the CVD subtype when the disease state is in its earliest forms. Accordingly, opportunity exists for the misclassification of risk. For example, a patient with non-obstructive coronary artery disease may still sustain a myocardial infarction due to plaque instability, which would not be found on conventional risk assessment.

The promise of artificial intelligence (AI) and machine learning in carrying out early detection of CVD is thrilling, but many issues remain with this as well. AI algorithms trained using electrocardiograms (ECGs) or medical imaging can detect very subtle patterns that point to disease, but these models are often limited by existing biases from the lack of representation of diverse groups in the training samples. The "black box" nature of some AI models renders them a poor fit for clinical use, as physicians may not be comfortable basing their clinical decisions on outcomes that provide no explanation of their prediction. There are also technical challenges in integrating AI tools into existing electronic health records (EHRs) because of issues related to format and systemic issues with quality that vary among systems.

In addition, economic and logistical barriers impede the widespread application of detection at an early stage. Indeed, advanced diagnostic tools (e.g., cardiac MRI, high-sensitivity biomarker assays) are often costly and present a burden especially in low-resource contexts. Furthermore, even if it could be established that such technologies were efficacious, their application in asymptomatic populations raises cost-effectiveness and overdiagnosis concerns. The ethical aspect of using early detection in the context of CVD should not be underestimated. Follow up tests could create unease for individuals who could be identified as being at high risk of CVD without binary clinical evidence to justify if these potential burdens are worth the investigation.

Genetic and epigenetic science has opened the door to new opportunities in early detection of CVD, but also presents new challenges. Polygenic risk scores (PRS) can be used to estimate an individual's genetic risk for CVD, but the predictive power differs greatly across ethnicities. Further, epigenetic modifications impacted by environmental and lifestyle factors introduce a new layer of complexity that limits our ability to convert genetic risk factors into clinically actionable insights. The interaction between genes and the environment means we need to start developing richer models that can use multiple data types for risk stratification.

Implementation challenges further complicate the progress in early detection. For instance, proteomic profiling and lipid particle characterization are advances in technology that have not developed practical clinical applications with standardized operating protocols. Applicability in practice is limited because approval processes have also slowed the application of detection based technologies. Moreover, uncertainty has arisen among clinicians on timelines to implement pharmacological treatment based on early detection findings as patients often first prefer lifestyle modifications to account for health issues. In addition, the overlap in pathophysiological factors of CVD with diabetes and kidney disease, as an example, all share similar molecular pathways and risk factors which complicates the diagnostic process. Beyond clinical and technological hurdles, patients remain reluctant to undergo testing for asymptomatic diseases, above and beyond symptom-based evaluations and physicians are working within healthcare systems that have trouble addressing cost-effectiveness when thinking of widespread screening initiatives for detecting early diagnosis.

As noted above, the current early detection capability draws limitations from various fronts and encompasses complex considerations that require more sophisticated risk stratification, a standardized detection protocol and screening initiatives that account for various factors mentioned above.

HOW CHALLENGES ARE CONVENTIONALLY HANDLED

In light of the above limitations, new methods are being developed, which could change the way cardiovascular diseases are identified early. New biomarker research using proteomic and omics discovery has identified new molecular fingerprints of early vascular dysfunction, which may be used to detect disease earlier than structural changes appear detectable. The use of AI and machine learning in risk prediction modelling means risk assessment can be more complex, as multiple datasets can inform the estimates of risk, both normalised (e.g. electronic patient records) and non-normalised (imaging, genetics, etc.). This is becoming easier for clinicians as these computational methods will be enhanced in next-generation imaging, such as high-resolution PET/MRI that can detect not only non-calcified but vulnerable, inflamed plaques. The basic practice for asymptomatic patients includes risk factor assessment and optimally, risk stratification; using well established scores such as the Framingham Risk Score (along with the intermittent risk factor assessment) followed by selective measurements in higher-risk patients. The authors of this article questioned the use of traditional risk factors to identify patients as not enough at-risk patients are identified by such traditional methods. Current recommendations for assessing intermediate risk patients includes coronary artery calcium scoring, but the cohorts may not always demonstrate calcium, and there continues to be a group of at-risk individuals with exposed angina, non-calcified plaques or microvascular disease.

The challenge of biological variability is addressed in traditional clinical research by using repeated biomarker measurements and by evaluating multiple biomarkers, although each of these options increases costs and complexity. In the case of plaque progression, serial imaging is used, albeit infrequently, in research settings, but it is largely impractical due to radiation exposure and expense for regular clinical care.

With the rapid technological advancements in wearable devices and digital health platforms, new approaches for continuous cardiovascular monitoring outside of clinical encounters are being initiated. New smartwatch-ECG and blood pressure monitoring systems can detect atrial fibrillation and hypertension, as with wearable sensors designed to record heart rate and rhythm continuously. Digital health platforms allow for real-time risk assessment in individuals that have, and are likely to remain, in a high-risk population following hospital discharge. Most intriguingly, are the new precision prevention research approaches where genetic risk profiling is merged with tailored therapeutic options, including the exciting and novel anti-inflammatory regimens that block early atherogenic processes. Collectively, these reported solutions suggest a change from reactive cardiovascular care, in which the patient is confronted to adapt their lifestyle, to proactive cardiovascular care that could help pre-empt the need for lifestyle or therapeutic changes. Acceptance of these emerging paradigms does provide challenges to assure their applicability in the broader population and acceptance as standard care. Ultimately, it is likely that cardiovascular disease prevention will emerge from the marriage of traditional preventive approaches and new technologies that result in better-sensitive and accessible personalized early detection options.

A major advancement in AI-based cardiovascular disease detection is the establishment of explainable AI (XAI) models that can overcome the long-standing "black box" problem inherent in traditional machine learning models. These new algorithms make risk predictions but are also able to provide clinically interpretable rationales - describing the specific physiological parameters - for instance, nocturnal blood pressure patterns or stress-related changes in an ECG tracing - that contributed most to the risk prediction.

What is novel about these AI systems is that it achieves two things at once: (1) it enhances the Clinical Users' confidence in AI-based risk predictions, and (2) it enables more guided personalized strategies for the user to reduce risk based on the patient's most prominent risk factors, rather than a consensus-based clinical approach. In parallel, concerning improvements in imaging, the optimized convolutional neural networks have enabled extraction of diagnostic information from lower resolution imaging data and may help to provide earlier detection capabilities to healthcare scenarios with limited access to the high-end imaging devices (e.g. CT, MRI) utilized when examining the heart and vascular systems.

New federated learning methods are also an important development in AI applications for cardiovascular disease. The ability to train models between different institutions and patient populations, without needing to centralize sensitive health data is enabled by distributed learning; and allows data privacy to be maintained while also gathering information from a wider variety of demographic groups. Improvements in generalizability can be achieved naturally, since older models, often based on data from a single or few sites, may suffer from performance bias relative to minority populations. The field of AI is also working with new and unique multimodal fusion techniques that combine heterogenous data streams that are traditionally disconnected - for example, the use of retinal vascular analysis combined with vocal biomarker patterns to observe initial cardiovascular changes, which may not be realized when only looking at either of the two sources isolated. AI is uniquely positioned to identify patterns such that these early warning signals may be discretely related across data sources, that human clinicians may not appreciate may be important together; these approaches give potential to advance earlier and improve risk stratification.

INDIVIDUAL PROJECT -> SNP - rs6922269

Genetic variations represent subtle yet powerful determinants of human health, influencing susceptibility to a wide range of diseases. Among these variations, the rs6922269 single-nucleotide polymorphism (SNP) has emerged as a significant focus of scientific investigation due to its location near the MTHFR gene on chromosome 1 and its potential role in critical metabolic processes. This essay examines the biological significance of rs6922269, explores its associations with various diseases, and analyzes its distinct distribution patterns across global populations, while considering the broader implications for personalized medicine.

Located in close proximity to the MTHFR gene, which encodes a key enzyme in folate metabolism, the rs6922269 SNP occupies a strategic position in human genetics. Although this genetic variation does not directly alter the structure of the MTHFR enzyme, it appears to function as a regulatory element that influences the gene's expression levels. Scientists classify rs6922269 as an expression quantitative trait locus (eQTL), a designation that indicates its potential to modulate the activity of nearby genes. Through this regulatory mechanism, the SNP may indirectly affect several crucial biological processes, including homocysteine metabolism, DNA methylation patterns, and various metabolic pathways. These wide-ranging influences help explain why this seemingly minor genetic variation has been associated with diverse health outcomes across multiple studies.

Research into the clinical significance of rs6922269 has revealed several important disease associations, primarily through genome-wide association studies (GWAS). The most robust evidence connects this polymorphism to cardiovascular diseases, particularly coronary artery disease and myocardial infarction. These associations likely stem from the SNP's potential impact on homocysteine levels, as elevated homocysteine is a well-established risk factor for vascular damage and thrombotic events. Beyond cardiovascular health, rs6922269 has also been implicated in metabolic disorders such as type 2 diabetes and obesity-related conditions, where alterations in folate metabolism may contribute to impaired glucose regulation and lipid processing. The polymorphism's potential role in oncogenesis has been explored as well, with some studies suggesting tentative links to colorectal and breast cancer risk, possibly mediated through folate-dependent DNA synthesis and repair mechanisms. In the neurological domain, researchers have investigated possible connections between rs6922269 and neurodegenerative conditions including Alzheimer's disease, hypothesizing that homocysteine-induced neurotoxicity might represent a contributing factor. However, it is important to note that many of these associations demonstrate relatively modest effect sizes and require further validation through additional research.

The distribution of rs6922269 across human populations presents a fascinating pattern of genetic variation. Population genetic studies have documented significant differences in the frequency of the minor allele (G) among ethnic groups, with approximately 30% of individuals of European ancestry carrying this variant compared to about 20% in East Asian populations and only 10% in African populations. These disparities may reflect various evolutionary forces, including selective pressures related to dietary adaptations or genetic drift in isolated populations. The concept of linkage disequilibrium becomes particularly relevant when considering how rs6922269 interacts with neighboring genetic variants, as these relationships can vary substantially between populations and potentially influence disease risk profiles. Such population-specific genetic architectures underscore the importance of including diverse ethnic groups in genetic research to ensure comprehensive understanding of disease mechanisms and equitable development of precision medicine approaches.

In conclusion, the rs6922269 polymorphism serves as a compelling illustration of how subtle genetic variations can exert significant influence on human health through complex regulatory mechanisms. While much has been learned about its potential roles in disease pathogenesis and population genetics, numerous questions remain unanswered. Future research directions should prioritize elucidating the precise molecular mechanisms through which this SNP operates, investigating potential gene-environment interactions, and exploring its clinical utility in personalized prevention and treatment strategies. As our understanding of genetic variants like rs6922269 continues to deepen, so too does the promise of developing more targeted and effective approaches to disease management and health promotion. The study of such polymorphisms not only advances our fundamental knowledge of human biology but also brings us closer to realizing the full potential of genomic medicine in improving population health outcomes.

SOURCES:

1. DeGroat, W., Abdelhalim, H., Peker, E., Sheth, N., Narayanan, R., Saman Zeeshan, Liang, B. T., & Ahmed, Z. (2024). Multimodal AI/ML for discovering novel biomarkers and predicting disease using multi-omics profiles of patients with cardiovascular diseases. Scientific Reports, 14(1). https://doi.org/10.1038/s41598-024-78553-6
2. Heusch, G. (2024). Myocardial ischemia/reperfusion: Translational pathophysiology of ischemic heart disease. Med, 5(1), 10–31. https://doi.org/10.1016/j.medj.2023.12.007
3. Gaze, D. C. (2009). Ischemia Modified Albumin: A Novel Biomarker for the Detection of Cardiac Ischemia. Drug Metabolism and Pharmacokinetics, 24(4), 333–341. https://doi.org/10.2133/dmpk.24.333

4. Franco, M., Cooper, R. S., Bilal, U., & Fuster, V. (2011). Challenges and Opportunities for Cardiovascular Disease Prevention. The American Journal of Medicine, 124(2), 95–102. https://doi.org/10.1016/j.amjmed.2010.08.015
5. Liu, R., Wang, M., Tang, Z., Zhang, R., Li, N., Chen, Z., Yan, H., & Shi, Q. (2022). An artificial intelligence-based risk prediction model of myocardial infarction. BMC Bioinformatics, 23(1). https://doi.org/10.1186/s12859-022-04761-4
6. Jacquet, S., Yin, X., Sicard, P., Clark, J., Kanaganayagam, G. S., Mayr, M., & Marber, M. S. (2009). Identification of Cardiac Myosin-binding Protein C as a Candidate Biomarker of Myocardial Infarction by Proteomics Analysis. Molecular & Cellular Proteomics, 8(12), 2687–2699. https://doi.org/10.1074/mcp.m900176-mcp200
7. Upadhyay, R. K. (2015). Emerging Risk Biomarkers in Cardiovascular Diseases and Disorders. Journal of Lipids, 2015, 1–50. https://doi.org/10.1155/2015/971453
8. Srinivasan, S. M., & Sharma, V. (2025). Applications of AI in Cardiovascular Disease Detection — A Review of the Specific Ways in which AI Is Being Used to Detect and Diagnose Cardiovascular Diseases. 123–146. https://doi.org/10.1002/9781394278695.ch6
9. Du, X., Su, X., Zhang, W., Yi, S., Zhang, G., Jiang, S., Li, H., Li, S., & Xia, F. (2021). Progress, Opportunities, and Challenges of Troponin Analysis in the Early Diagnosis of Cardiovascular Diseases. Analytical Chemistry, 94(1), 442–463. https://doi.org/10.1021/acs.analchem.1c04476
10. Celermajer, D. S., Chow, C. K., Marijon, E., Anstey, N. M., & Woo, K. S. (2012). Cardiovascular Disease in the Developing World. Journal of the American College of Cardiology, 60(14), 1207–1216. https://doi.org/10.1016/j.jacc.2012.03.074
11. Hassan, P. L. (2024). Artificial Intelligence in Early Detection and Prediction of Cardiovascular Diseases. Non-Communicable Diseases and Care, 1(1), 1–12. https://www.researchcorridor.org/index.php/ncdc/article/view/372
12. Ogunpola, A., Saeed, F., Basurra, S., Albarrak, A. M., & Qasem, S. N. (2024). Machine Learning-Based Predictive Models for Detection of Cardiovascular Diseases. Diagnostics, 14(2), 144. https://doi.org/10.3390/diagnostics14020144
13. Hassan Shabeer, Hafiz Muhammad Ali Haider, Tamana Khatri, Nouman Khan, Adnan Ahmed Rafique, & Fakhar Anjam. (2025). The Role of Artificial Intelligence in Early Diagnosis and Management of Cardiovascular Diseases. Indus Journal of Bioscience Research, 3(2), 213–224. https://doi.org/10.70749/ijbr.v3i2.667
14. Wu, A. H. B. (2005). Markers for Early Detection of Cardiac Diseases. Scandinavian Journal of Clinicaland Laboratory Investigation, 65(sup240), 112–121. https://doi.org/10.1080/00365510500236283