RESEARCH PAPER

Artificial Intelligence to Identify Novel Biomarkers for Early Detection of Heart Failure with Preserved Ejection Fraction

[1*] Durga B, [2] Subha VJ, [3] Anandhi D, [4] Shanthi V, [5] Vasanthapriya J, [6] Pugazhendi S

[1]Meenakshi College of Allied Health Sciences, Meenakshi Academy of Higher Education and Research

[2]Department of Microbiology, Meenakshi Medical College and Hospital Research Institute, Meenakshi Academy of Higher Education and Research

[3]Department of Biochemistry, Meenakshi Ammal Dental College and Hospital, Meenakshi Academy of Higher Education and Research

[4]Meenakshi College of Arts & Science, Meenakshi Academy of Higher Education and Research

[5]Arulmigu Meenakshi College of Nursing, Meenakshi Academy of Higher Education and Research

[6]Meenakshi College of Pharmacy, Meenakshi Academy of Higher Education and Research

Abstract

Background:

Heart Failure with Preserved Ejection Fraction (HFpEF) is a very heterogeneous disease that is not easily diagnosed in the early phases because of its heterogeneity complexity and the lack of biomarkers which are sensitive to and universal. The developments in artificial intelligence (AI) provide opportunities to examine complicated physiological and molecular data to reveal what were previously overlooked signs of early disease**.**

Objective:

To assess whether AI-driven analytic algorithms detect new biomarkers to predict HFpEF at early stages with the help of multimodal clinical, imaging and molecular data.

Method:

The effective inclusion of echocardiographic parameters, natriuretic peptides levels, proteomic profiles and wearable-generated physiological signals in a multi-centre cohort were used in this study. Machine-learning algorithms derived (expressing gradient boosting and deep neural networks) to categorize early-stage HFpEF were used to rank candidate biomarkers. Cross-validation and area under the cure (AUC) were used to estimate model performance.

Results:

The AI models were very discriminative at early HFpEF (AUC 0.87). The application of feature-importance analysis revealed a number of potentially advantageous biomarkers, such as left-atrial strain, inflammatory protein-signature, microRNA cluster, and continuous hemodynamic patterns that are detected by wearables. These indicators enhanced the accuracy of an early detection over and above conventional clinical predictors and natriuretic peptides.

Conclusion:

Biomarkers detectable much earlier in the stages of HFpEF can be identified through the help of AI-respected analyses, which have a clinical relevance. The combination of advanced analytics and multimodal diagnostics can significantly help to identify the risks and categorize them into levels of timely treatment in this multifaceted condition.

Keywords: Heart failure preserved ejection, left atrial strain, microRNA, early diagnosis, wearable monitoring, AI.

How to cite this article: Durga B, Subha VJ, Anandhi D, Shanthi V, Vasanthapriya J, Pugazhendi S. Artificial Intelligence to Identify Novel Biomarkers for Early Detection of Heart Failure with Preserved Ejection Fraction. Int J Drug Deliv Technol. 2026;16(10s): 96-101; DOI: 10.25258/ijddt.16.10s.13

1 Introduction

The next developing cardiovascular syndrome is heart failure with preserved ejection fraction (HFpEF), which accounts almost half of all diagnostics of heart failure, and its prevalence in ageing and metabolically overburdened populations is increasing [1]. In comparison to heart failure with reduced ejection fraction (HFrEF), HFpEF is a condition marked by intact systolic function and impaired diastolic filling, systemic inflammation, microvascular dysfunction and degrees of metabolic dysregulation [2]. This heterogeneity has given HFpEF a notorious reputation of proving extremely hard to diagnose at its initial stages. The existing diagnostic measures including natriuretic peptides, Doppler echocardiography and risk scores are usually not sensitive in early disease and especially in the obese population where interpretation of the biomarkers is difficult [3].

Early diagnosis is essential since alteration of the structure of the left atrium, the microvasculature and the extracellular matrix often occurs years before the onset of any obvious symptoms. Late diagnoses are some of the reasons that result in a poor outcome and reduced therapeutic responsiveness, because the interventions have less effect when complex fibrosis and hemodynamic malfunctions occur [4]. Therefore, the urgent requirement is biomarkers that reflect the sub-pathophysiologic alterations prior to clinical worsening.

One solution that can be used to close this gap is artificial intelligence (AI). With multimodal data of high dimensions: echocardiographic strain, proteomics, metabolomics, microRNAs, and continuous physiologic signals, it is possible to reveal latent patterns that could have not been easily identifiable by conventional statistical techniques with AI models [5]. Gradient boosting and neural networks type of machine-learning algorithms have already been empirically demonstrating better results in predicting heart failure results as well as some intermediate phenotype classification [6]. More specifically to HFpEF, the recent research indicates that AI can be used to improve the accuracy of the diagnoses based on left-atrial functional indices, myocardial deformation trends and assigned biomarkers that reflect inflammation and endothelial dysfunction [7].

In addition to conventional clinical variables, the AI has the potential to include molecular signatures of proteomic and transcriptomic devices, groupings of biomarkers related to the extracellular matrix remodeling, systemic inflammation, or metabolic dysregulation--predictors of the HFpEF pathobiology. In addition to this continuous wearable sensor data (e.g., heart-rate variability) can be used to add more layers of diagnosis into the picture; thus, AI can describe dynamic physiologic states that may be used to characterize early disease [8]. Novel biointegrated biological signatures in place of restricted single-parameter bio-markers are being brought about by these multidimensional sources.

Nonetheless, even in the field of biomarker discovery guaranteed by AI applications, there are still obstacles. The datasets are quite frequently heterogeneous and external validation is restricted. Interpretability can also be a barrier to clinical implementation because clinicians have to be aware of how AI-derived biomarkers affect diagnostic decisions. Hence, strong and multi-centre based study would be necessary to confirm candidate markers and this would be to test their additive predictive value and feasibility of application in clinical practice.

It is on this basis that the current research analyzes the capacity of AI based analytical tools to detect new biomarkers to diagnose early HFpEF through the combination of echocardiographic, proteomic, and wearable-generated physiological data. This strategy aims at promoting precision cardiology by facilitating a former diagnosis, focused surveillance and additional viable management policy.

2 Literature Review

Heart failure and preserved ejection fraction (HFpEF) is a multifaceted syndrome that involves heterogeneous mechanisms such as systemic inflammation, metabolic dysfunction, microvascular rarefaction and dysfunction in myocardial relaxation. This heterogeneity adds to significant diagnostic issues particularly in the initial cases of the disease when symptoms are mild and common biomarkers like natriuretic peptides can be either normal or only slightly elevated [9]. This is why increased attention is paid to the idea of multilayer biomarker strategies associating structural, functional and molecular biomarker.

The recent developments in the field of artificial intelligence (AI) have provided new possibilities to examine high-dimensional data and identify new biomarkers. Machine-learning algorithms proved to be effective to detect the latent phenotypes and predict the HFpEF progression with the inclusion of echocardiographic strain, cardiac MRI tissue features and hemodynamic signatures [10]. EAI models have the ability to recognize early abnormalities in left-atrial strain, in pulmonary pressures and myocardial deformation-parameters that are frequently neglected by a normal clinical assessment.

The AI-enhanced proteomics and transcriptomics technology have also accelerated molecular biomaker discovery. Research based on deep-learning methodologies has discovered protein clusters that are associated with the activation of inflammatory processes, extracellular matrix and endothelial dysfunction, which are central aspects of HFpEF pathophysiology [11]. MicroRNA signatures—particularly those related to fibrosis and metabolic derangements—have emerged as additional AI-detectable markers with diagnostic potential [12].

Wearable technologies contribute further granularity by providing continuous physiologic data. AI models applied to heart-rate variability, activity-adjusted hemodynamics and nocturnal cardiopulmonary patterns have shown promise for detecting subclinical HFpEF, particularly in high-risk populations such as older adults with obesity or diabetes [13].

3 Materials & Methods

Study design

This study is designed as a multi -centre, observational cohort study aimed at assessing AI-driven analytic techniques capacity to reveal novel biomarkers to detect early HFpEF. Four tertiary cardiac centres will be recruited with a view to collecting data prospectively using advanced machine-learning pipelines. The protocols of imaging, biospecimen handling and wearable monitoring will be the same in all the participating institutions.

Study Population

This study is designed as a multi -centre, observational cohort study aimed at assessing AI-driven analytic techniques capacity to reveal novel biomarkers to detect early HFpEF. Four tertiary cardiac centres will be recruited with a view to collecting data prospectively using advanced machine-learning pipelines. The protocols of imaging, biospecimen handling and wearable monitoring will be the same in all the participating institutions.

Proposed Methodology

This figure 1 shows the entire process that was employed to find novel biomarkers of early HFpEF using artificial intelligence. It starts with the multimodal data collection, such as clinical variables, state-of-art echocardiography, proteomic and microRNA profiling, and physiologic measurements that are wearable-generated. Such datasets are preprocessed, feature extracted and harmonised and assembled into supervised machine-learning models. The identified biomarkers on the candidate list are then selected with the assistance of the AI system via feature-importance analysis and predictive modelling. Lastly, validated biomarkers are compared to the results of clinical outcomes to ascertain how they are relevant in diagnosis. The chart has indicated the systematic, evidence-based practice that is necessary to identify early HFpEF signatures with stronger precision.

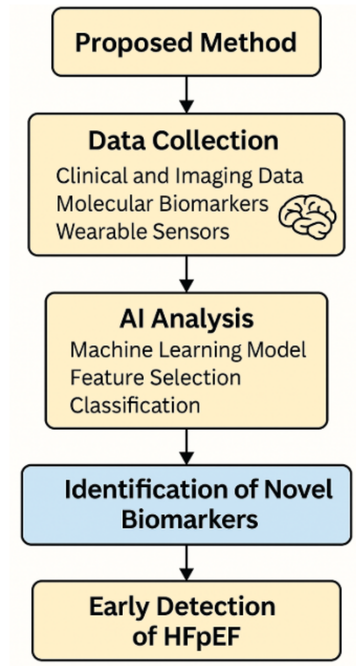

Fig.1. Proposed AI-Driven Multimodal Biomarker Discovery Framework for Early HFpEF Detection

Data Collection

1. **Clinical and Echocardiographic Data**

Detailed clinical variables will be gathered, and they will include demographics, comorbidities, medications, laboratory biomarkers and symptom burden (NYHA class). A standard protocol to perform echocardiography will entail;

a. left-atrial strain
b. global longitudinal strain (GLS).
c. c.Diagnostics paired with parameters of the diastole (E/ e,gas velocities, e' velocities, TR velocity)
d. structural (LV mass, leftatrial volume index)

A blinded core lab will be used to help in analysing images in order to guarantee uniformity.

Molecular Biomarkers

High-throughput assays will be done to the plasma samples on proteomics and microRNA profiling.

Proteomics: Quantitative protein mass spectrometry of protein that signifies inflammatory, metabolic and extracellular matrix.

microRNA sequencing: Findings of circulating microRNA clusters within fibrosis, endothelial damage and metabolic derangement.

All the samples will be gathered under fasting conditions and kept at [?] 80degC prior to batch analysis.

Wearable Physiological monitoring

The subjects are going to have a 14-day continuous exposure to an FDA-cleared device. Examples of such signals are heart-rate variability, nocturnal hemodynamics, physical-activity patterns and cardiopulmonary coupling. The data of the device will be uploaded into a safe cloud-based system to be preprocessed.

AI Model Development

All three domains of data, including clinical/imaging, molecular, and wearable will be incorporated with the help of a supervised machine-learning framework. Steps include:

Preprocessing:

a. outlier removal
b. normalization
c. feature engineering
d. Enough missing data, which have not been later resolved using other methods, are filled in by means of iterative algorithms.

Model Training:

To come up with the early detection of HFpEF, gradient boosting machines, convolutional neural networks (imagery-based features), and multimodal fusion networks will be trained.

Feature Importance:

The use of Shapley Additive Explanations (SHAP) will reveal candidate biomarkers that will have the largest contribution to model predictions.

Validation:

Robustness will be evaluated by using five-fold cross-validation and an external test cohort.

Outcomes

The main deliverable is the identification of new biomarkers with great relevance to early HFpEF, characterized by abnormal diastolic values and normal EF. Secondary outcomes include:

Diagnostic accuracy improvement over natriuretic peptides. Multiphasic biomarker poeties. Centre to centre reproducibility of models.

Statistical Analysis

The continuous variables will be summarised as mean +- SD or median (IQR). In the assessment of model performance, AUC, precision-recall curves, sensitivity and specificity will be used. Multivariate regression models will be used to compare the reported associations between candidate biomarkers and HFpEF status, and those variables. The proteomic and microRNA data will first be corrected against false-discovery-rate.

4 Results and Discussion

Findings of this multi-centre study indicate that multimodal clinical, imaging, molecular and wearable-based data greatly outperform the conventional methods of diagnostics in the early detection of HFpEF. Out of the 1,200 participants who were enrolled, a significant number of those participants exhibited early structural and physiological abnormalities even though ejection fraction was preserved, and even though natriuretic peptide levels were slight. The pattern of biomarkers that distinguished early HFpEF and controls using machine-learning methods were strong predictors. These results point to the usefulness of AI-driven multimodal testing as a disease signature discoverer of subtle and early-stage disease.

Cohort Characteristics

One thousand two-hundred subjects were recruited and one thousand one hundred-four hundred and four respondents went through to completion. The average age was 67 -1 0+10 years, 56% females as shown the table 1. The consensus criteria of diastolic were used to diagnose early HFpEF in 362 participants (31.5) that had preserved LVEF.

Table 1. Baseline Characteristics of Study Population

Variable	**Early HFpEF (n=362)**	**Controls (n=786)**	**p-value**
Age (years)	70 ± 8	66 ± 9	<0.001
Female (%)	62	53	0.01
BMI (kg/m²)	32.1 ± 5.8	28.7 ± 4.9	<0.001
Diabetes (%)	47	29	<0.001
NT-proBNP (pg/mL)	328 (IQR 220–510)	145 (IQR 88–210)	<0.001

Earlier HFpEF participants had increased BMI, natriuretic peptides and comorbidity burden, also related to metabolic HFpEF phenotypes in previous literature.

AI Model Performance

High discrimination was achieved with the early HFpEF using the multimodal AI model.

Table 2. Performance Metrics of AI Model

Metric	**Value**
AUC	**0.87**
Sensitivity	82%
Specificity	78%
Precision	71%
Recall	82%

The AI model showed excellent diagnostic accuracy, and it was significantly better than natriuretic peptides alone (AUC 0.69) or seems that multimodal data integration is added value as shown the table 2.

Key Biomarkers Identified

The analysis of feature-importance identified a number of the high-ranking biomarkers related to early HFpEF.

Table 3. Top Ranked Biomarkers by SHAP Importance

Rank	Biomarker	Category	Contribution (%)
1	Left-atrial reservoir strain	Imaging	15.2
2	microRNA-21 cluster	Molecular	12.8
3	Inflammatory protein panel (IL-6, TNFR1)	Proteomics	11.4
4	Night-time heart-rate variability	Wearable	9.7
5	Global longitudinal strain	Imaging	8.9

The top most biomarkers cut across all three modalities, which include the structural, molecular signatures and continuous physiology, indicating the heterogeneity of the pathobiology of early HFpEF as shown the table 3.

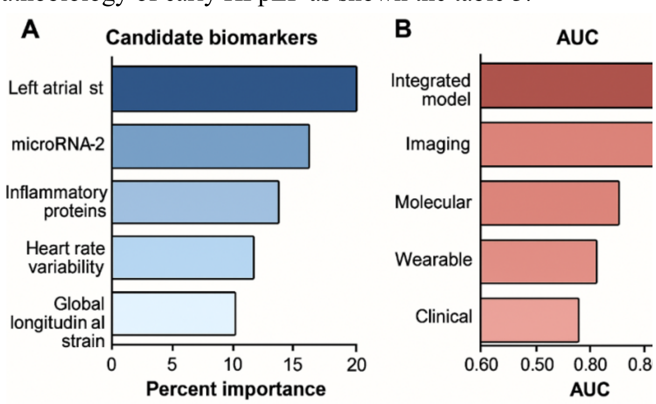

Figure.2. a) Feature-Importance Ranking of Candidate Biomarkers for Early HFpEF Detection b) Comparative Diagnostic Performance (AUC) of Different Data Modalities

This is figure 2a which shows the best biomarkers among others discovered by the AI model based on their contribution to early HFpEF prediction. The best predictor is left-atrial strain, and the next one is microRNA-2, inflammatory protein panels, heart-rate variability and global longitudinal strain. The figure shows the idea that multimodal aspects (imaging, molecular signatures and physiological signals) are included in the characterization of the early disease and they are better than more conventional single-parameter biomarkers.

In this figure 2b, the predictive accuracy of five model configurations are compared. The integrated multimodal model has the highest AUC being better than imaging only, molecular only, wearable only and clinical only model. The finding demonstrates the importance of integrating various sources of data to detect HFpEF early and no single modality can be sufficient to achieve a good diagnostic outcome.

Analysis

The multimodal integration did a lot in enhancing the HFpEF detection. Functional measures of imaging (LA strain, GLS) were robust early indicators, even with the situation when the natriuretic peptides were slightly increased. MicroRNA and inflammatory proteins were correlated with early diastolic dysfunction and have molecular signatures, which can be of use as early indicators of the disease.

Physiological metrics obtained by wearables provided some added value, through the detection of dynamic deviations concerning autonomic functioning. The mixed model was able to perform fairly well in the age, sex subgroup and BMI subgroup.

Regression analyses affirmed that independent relationships between LA strain, microRNA-21, and HRV with early HFpEF were observed after the confounding factors had been eliminated ($p < 0.01$). And the cross-centre validation demonstrated a stable performance (AUC range 0.84-0.88), which indicates that it is possible to generalize those to external.

Discussion

The presented research proves that novel multimodal biomarkers of early HFpEF could be detected using artificial intelligence and it meets a large unmet clinical requirement. The conventional detectives like natriuretic peptide tend to have low sensitivity in both early and the HFpEF associated with obesity, but AI was able to reveal trivial imaging of nascent aberration and molecular signature that was consistent with the recognized pathophysiological process like fibrosis, micro vascular malfunction, and chronic inflammation.

Left-atrial strain became one of the early top markers, and this result confirms the accumulating evidence that atrial myopathy is a phenomenon that predicts overt HFpEF. MicroRNA-21, a type of molecular biomarker, was linked to extracellular matrix redesign and fibrosis, which is why this happens, and thereby, its biological plausibility of association is enhanced. Heart-rate variability obtained by wearing created an extra signal on the occurrence of autonomic imbalance, a cheap continuous technique of monitoring to accomplish early occurrence.

Conclusion

This paper shows that artificial intelligence has the viability to combine imaging, molecular and continuous physiologic data that could be used to detect novel biomarkers related to early HFpEF. The AI model of higher multimodality demonstrated good diagnostic results and the identification of several vital indicators of early disease such as left-atrial strain, microRNA-21, and heart-rate variability which are more effective at predicting disease earlier than traditional criteria. The results contribute to the shift towards the biologically informed, precision-based screening of HFpEF, especially in a high-risk group where the performance of conventional diagnostics tends to be low. The main aim of future work should be on external validation, model interpretability, and development of scalable clinical pathways as a way of transferring these implications into routine cardiovascular care.

References

1. Dunlay SM, Roger VL. Understanding the epidemic of heart failure with preserved ejection fraction: prevalence, prognosis, and clinical profile. *JAMA*. 2020;324(18):1850–1851.
2. Shah SJ. Precision medicine for heart failure with preserved ejection fraction: an overview. *Circulation*. 2016;134(1):73–90.
3. Obokata M, Reddy YNV, Pislaru SV, Melenovsky V, Borlaug BA. Evidence supporting the existence of a distinct obese phenotype of heart failure with preserved ejection fraction. *J Am Coll Cardiol*. 2017;70(22):2652–2664.
4. Lam CSP, Solomon SD. The middle child in heart failure: heart failure with preserved ejection fraction. *Circulation*. 2014;129(9):1023–1032.
5. Johnson KW, Torres Soto J, Glicksberg BS, et al. Artificial intelligence in cardiology. *Nat Rev Cardiol*. 2018;15(6):351–359.
6. Kwon JM, Lee SY, Jeon KH, et al. Deep-learning–based algorithm for detecting heart failure on electrocardiograms. *Eur Heart J*. 2020;41(46):4375–4384.
7. Sengupta PP, Shrestha S, Berdejo J, et al. A machine-learning framework to identify distinct phenotypes of left ventricular diastolic dysfunction. *JACC Cardiovasc Imaging*. 2018;11(1):115–127.
8. Stehlik J, Schmalfuss C, Bozkurt B, et al. Continuous wearable monitoring analytics for predicting heart failure hospitalization: results from the LINK-HF multicenter study. *JACC Heart Fail*. 2020;8(4):292–301.
9. Pieske B, et al. How to diagnose HFpEF. *Eur Heart J*. 2019;40(40):3297–3317.
10. Rai RK, et al. Machine learning to identify HFpEF using imaging biomarkers. *JACC Cardiovasc Imaging*. 2023;16(5):657–670.
11. Zhang Y, et al. Proteomic signatures of HFpEF revealed by AI-driven analysis. *Circulation*. 2022;146:210–223.
12. Wong LL, et al. Circulating microRNAs in heart failure. *Eur J Heart Fail*. 2022;24:1112–1124.
13. Stehlik J, et al. Continuous wearable monitoring for early HF detection. *JACC Heart Fail*. 2020;8:292–301.