Review

ANNALS OF MEDICINE & SURGERY

OPEN

Machine learning in heart failure diagnosis, prediction, and prognosis: review

Muhammad Saqib, MBBS[a,*], Prinka Perswani, MD[e], Abraar Muneem, MD[f], Hassan Mumtaz, MBBS, MSPH, MRSPH[g], Fnu Neha, MBBS[b], Saiyad Ali, MBBS[c], Shehroze Tabassum, MBBS[d]

Abstract

Globally, cardiovascular diseases take the lives of over 17 million people each year, mostly through myocardial infarction, or MI, and heart failure (HF). This comprehensive literature review examines various aspects related to the diagnosis, prediction, and prognosis of HF in the context of machine learning (ML). The review covers an array of topics, including the diagnosis of HF with preserved ejection fraction (HFpEF) and the identification of high-risk patients with HF with reduced ejection fraction (HFrEF). The prediction of mortality in different HF populations using different ML approaches is explored, encompassing patients in the ICU, and HFpEF patients using biomarkers and gene expression. The review also delves into the prediction of mortality and hospitalization rates in HF patients with mid-range ejection fraction (HFmrEF) using ML methods. The findings highlight the significance of a multidimensional approach that encompasses clinical evaluation, laboratory assessments, and comprehensive research to improve our understanding and management of HF. Promising predictive models incorporating biomarkers, gene expression, and consideration of epigenetics demonstrate potential in estimating mortality and identifying high-risk HFpEF patients. This literature review serves as a valuable resource for researchers, clinicians, and healthcare professionals seeking a comprehensive and updated understanding of the role of ML diagnosis, prediction, and prognosis of HF across different subtypes and patient populations.

Keywords: biomarkers, gene expression, heart failure, patient readmission, prognosis

Introduction

Ventricular systolic or diastolic dysfunction resulting from structural or functional heart abnormalities characterizes heart failure (HF)[1]. As the final stage of diverse heart diseases, HF accounts for one-third of worldwide deaths from cardiovascular disease, according to the American College of Cardiology. In the United States, over 5 million people suffer from HF, with 550 000 new cases diagnosed annually[2,3]. In China, ~8.9 million people have HF, with a 1.3% prevalence rate for those over 35 years old[4]. The global mortality rates and increasing prevalence of HF make it a significant public health concern, with annual costs estimated at $29 billion due to high hospitalization rates and unsatisfactory prognoses[5]. Predicting mortality can assist doctors in creating appropriate treatment plans, preventing worsening conditions, reducing medical expenses, and enhancing quality of life.

HIGHLIGHTS

- Convolutional neural networks excel in heart failure with preserved ejection fraction patient differentiation.
- Long short-term memory shows promise in heart failure with reduced ejection fraction risk assessment.
- Machine learning surpasses traditional methods in heart failure mortality prediction, especially support vector machines in lower-middle income countries.
- XGBoost excels in ICU mortality prognosis.
- HFmeRisk offers early assessment for heart failure with preserved ejection fraction integrating epigenetics.

HF is linked to common indicators like difficulty breathing, swelling in the legs, and feeling tired, alongside physical signs such as crackling sounds in the lungs during examination and increased pressure in the jugular veins[6]. Currently, natriuretic peptides and common HF signs and symptoms are used to diagnose HF with preserved ejection fraction (HfpEF), which is then classified according to the left ventricular ejection fraction (LVEF)[7]. Clinical investigations have varied the cut-off LVEF for HFpEF between 40%, 45%, and 50%. According to the most recent recommendations[7], HFpEF is identified when a patient exhibits symptoms and indications of HF with an LVEF below 50% and indications for higher vascular volume (a rise in natriuretic peptides, for example), myocardial abnormalities, or both.

HFpEF is a complex clinical illness due to its high genetic heterogeneity and varied presentations[8,9], as well as a possibly

[a]Khyber Medical College, Peshawar, [b]Jinnah Sindh Medical University, Karachi, [c]Saidu Medical College, Swat, [d]King Edward Medical University, Lahore, Pakistan, [e]University of California Riverside, Riverside, [f]College of Medicine, The Pennsylvania State University, Hershey, United States and [g]BPP University, United Kingdom

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

*Corresponding author. Address: Khyber Medical College, Peshawar, Pakistan, Road no. 2, University of Peshawar, University Road, Peshawar 25120, Pakistan. E-mail: muhammadsaqib.drkmc@gmail.com (M. Saqib).

Annals of Medicine & Surgery (2024) 86:3615–3623

Received 9 February 2024; Accepted 24 April 2024

Published online 6 May 2024

http://dx.doi.org/10.1097/MS9.0000000000002138

nonlinear link between genes and clinical results. Thus, traditional linear generalized models [such as logistic regression (LR)] are subpar for risk forecasting. It is recommended that advanced statistical approaches and ML methods be researched for HFpEF prediction because they have the potential to improve classification performance over traditional statistical tools by taking into account nonlinear impacts of variables to arrive at an accurate prediction[9].

HF can stem from various underlying conditions, with cardiac amyloidosis being a notable contributor. It is important to rule out amyloidosis during the differential diagnosis of HF, especially in older patients without conventional risk factors for HFpEF (hypertension, obesity, and type 2 diabetes mellitus)[10,11]. Dyspnea, edema, angina, and syncope are the most typical signs of heart disease. Peripheral neuropathy, numbness, neuropathic pain, and loss of muscular strength in the lower limbs are examples of noncardiac symptoms that may raise concern. Autonomic neuropathy has been linked to gastrointestinal symptoms as nausea, vomiting, diarrhea, and weight loss. Sexual problems, low blood pressure when standing, and bladder problems caused by nerve damage are all autonomic symptoms[12]. Historically, a biopsy was the 'gold standard' for diagnosing amyloidosis[12].

HF can also result from cardiomyopathies. The most prevalent kind of familial cardiac disease is hypertrophic cardiomyopathy. Hypertrophy of the left ventricular (LV) muscle that is unrelated to increasing afterload characterizes hypertrophic cardiomyopathy. There is a hypertrophy, disorganization, and fibrosis of myocytes, as seen on a histopathology slide. Common signs include syncope, ventricular tachycardia, and orthopnea (dyspnea at rest)[13,14].

However, identifying HFpEF is still difficult. Elevated pulmonary capillary wedge pressure at rest 15 mmHg or during activity 25 mmHg by right catheterization is the gold standard test for verifying HFpEF. The diagnostic criteria for HFpEF in the current recommendations are not without some degree of uncertainty. Natriuretic peptide and echocardiographic data are used to diagnose HFpEF, although both have limited sensitivities[15]. For instance, there may be only limited use in using natriuretic peptides to assess HFpEF[16]. Natriuretic peptides were within normal limits in a sizable proportion of HFpEF patients despite the presence of clinical, echocardiographic, and hemodynamic markers of HF[17].

Diagnosing HFpEF can be difficult since the symptoms are ambiguous and may be caused by illnesses other than HF, such as anemia, chronic lung disease, and chronic kidney disease[18]. Importantly, in a patient presenting with HF symptoms, a diagnosis of HFpEF can be made with only an EF > 50% and signs of elevated LV filling pressure (elevated E/e' ratio, raised left atrial volume, higher BNP or NT-proBNP, or increased invasive LV filling pressure). The lack of 'diastolic dysfunction' on echocardiography does not rule out the diagnosis of HFpEF if there is other objective evidence of high LV filling pressure at rest or with activity. This is because diastolic function classification can be relatively varied and subjective.

Recently, the H2FPEF and HFAPEFF scores[19] have been developed to address the diagnostic conundrum of HFpEF. Age, BMI, AF, and hypertension are the four clinical components, whereas E/e′ and right heart valve pressure are the two echocardiographic components that make up the H2FPEF score. Minor and major criteria within the practical (E/e′, e′, regurgitated tricuspid velocity, global vertical strain), structural (left atrial volume score and measures suggesting LV hypertrophy), and natriuretic peptides domains make up the HFAPEFF score. A diagnosis of HFpEF can be made with an H2FPEF score of 6 or an HFA-PEFF score of 5. The authors recommend invasive hemodynamic assessment, preferably with exercise, or exercise echocardiography for patients with an H2FPEF score of 2–5 or an HFAPEFF of 2–4 points[20,21].

The management of HF poses significant financial burdens in high-income nations[22]. While traditional diagnostic modalities have provided invaluable insights into HF pathophysiology, their limitations in capturing the multifaceted nature of the disease underscore the need for innovative approaches. Recent advancements in diagnostic scoring systems, exemplified by H2FPEF and HFAPEFF, offer promising avenues for addressing diagnostic uncertainties and optimizing patient care. In this context, ML emerges as a transformative tool capable of harnessing the wealth of clinical data to facilitate more accurate risk prediction, prognostication, and treatment optimization in HF. By leveraging complex algorithms and computational techniques, ML methodologies hold the potential to unravel intricate patterns within large datasets, thereby empowering clinicians with actionable insights for personalized patient care[23]. A range of methods has been employed for data collection and analysis of HF, including the application of ML classifiers to predict patient survival, implementation of supervised deep learning (DL), and ML algorithms[24].

In this holistic review of the literature, we aim to shed light on the latest updates in an intuitive fashion on the use of ML for HF diagnosis, prediction and prognosis to give an updated overview to practicing care-givers and readers alike. A central illustrative figure depicts the gist of this review shown in Figure 1.

Heart failure diagnosis

Artificial intelligence (AI) is a technology that can perform tasks that typically require human intelligence. In the field of cardiovascular medicine, AI is being used more and more to transform the way we diagnose, treat, predict risks, provide clinical care to patients, and discover new drugs[25]. Unterhuber *et al.*[26] utilized

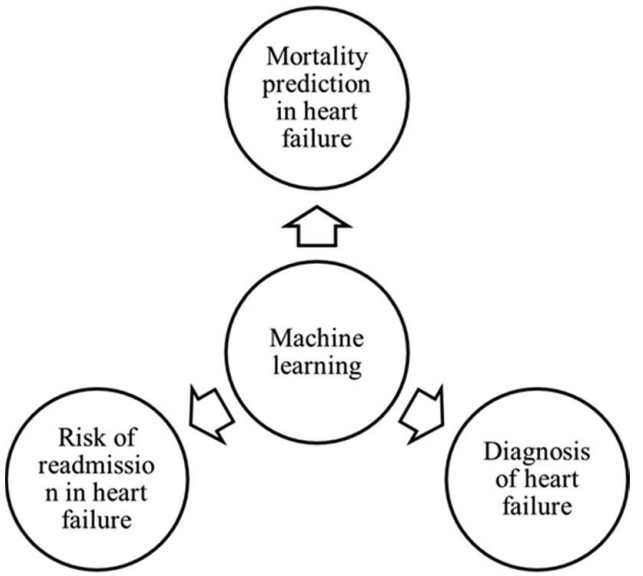

Figure 1. Machine-learning in heart failure diagnosis and prediction of readmission and mortality risks.

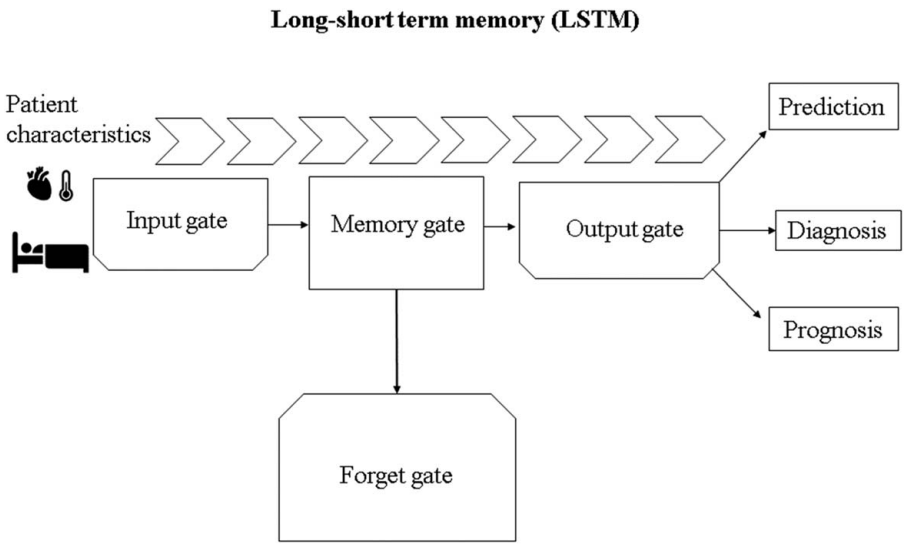

Figure 2. Long-short-term memory is a form of neural network that receives data on patient characteristics through the input gate. The memory gate is only a conveyor belt. The input gate selectively feeds data into the memory gate as needed. The forget gate deletes data that does not need to be remembered. The output gate leads to the needed output, which could be prediction, diagnosis, or prognosis.

a convolutional neural network (CNN) model to distinguish between individuals with HFpEF and controls. There were two patient cohorts in the trial. There were 1884 individuals in the derivation cohort who had exertional dyspnea or similar, preserved ejection fraction (50%) and clinical signs suggesting for coronary artery disease. The ECGs were segmented, providing 77 558 samples in total. The European Society of Cardiology (ESC) criteria categorize HFpEF and control patients using CNN. An external cohort of 203 participants from a prospective HF screening program served as the CNN's validation cohort. The CNN demonstrated a high discriminatory ability, achieving an AUC of 0.92 on the blinded test set, with a sensitivity of 0.98 and specificity of 0.63[27]. These findings demonstrated the first DL-enabled CNN for detecting patients with HFpEF based on ESC criteria, using NT-proBNP values in the diagnosis algorithm among at-risk patients. The CNN's appropriateness was evaluated on an external validation cohort of individuals at risk of developing HF, and the screening performance was essential.

These findings also highlight the potential of CNN-based ML models to aid in HFpEF diagnosis and management, offering valuable insights to clinicians for optimized patient care. Further research and validation are warranted for their practical integration. In another study, Wang *et al.*[27] focused on utilizing DL with long short-term memory to identify high-risk HFrEF patients using a large US nationwide commercial insurance dataset. Long short-term memory, a type of recurrent neural network, is commonly used in DL for sequential data analysis. For example, it can predict HF progression by analyzing a sequence of patient vital signs and medical measurements as schematically presented in Figure 2[28].

Heart failure prediction

Mortality in heart failure with reduced ejection fraction (HFrEF)

Tohyama *et al.*[28] emphasized the effectiveness of ML in predicting prognosis for HF patients using the Japanese Administrative Claims Database (ACD). The ML approach outperformed conventional risk models, leading the authors to develop a new prediction model called SMART-HF. By combining key variables identified through ML analysis, SMART-HF demonstrated equivalent or superior performance while requiring only a small number of easily accessible variables. These variables can be assessed through a brief interview, even by nonhealthcare providers, enhancing the model's usability. Among the ML algorithms considered, the voting classifier algorithm emerged as the most effective for predicting mortality in HF patients. This algorithm aggregates predictions from multiple experts using different methods and determines the outcome based on the majority. By leveraging the voting classifier algorithm, HF specialists can access more accurate predictions, facilitating decision-making in patient care. The permutation feature importance technique proved valuable in understanding the significance of different predictors, including mortality in HF. This technique involves shuffling the values of a single feature and observing its impact on prediction accuracy. If a feature is crucial, shuffling its values significantly reduces accuracy. The voting classifier algorithm is schematically described in Figure 3. By utilizing this technique, HF specialists can identify the most influential factors in predicting mortality, enabling them to prioritize essential aspects of patient management. Tohyama *et al.*[28] reported that evaluating the models' effectiveness in ranking HF patients based on mortality risk, metrics such as C-statistics, including AUC-ROC, were utilized. A higher C-statistic indicates a better ability to correctly order patients in terms of their predicted mortality probabilities. Although the differences in C-statistics between the models were marginal, SMART-HF offered distinct advantages compared to the other models that were studied. The use of ACD-based prediction models can also facilitate ML-based modeling research in various diseases in the future[29].

ML algorithms have also been utilized by others, for instance, Mpanya *et al.*[30] used ML to predict all-cause mortality in HF patients in lower-middle income countries. They trained six

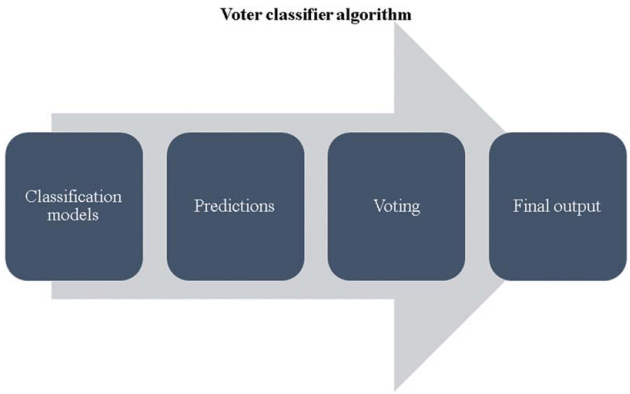

Figure 3. The voter classifier algorithm takes input from multiple classification models made by experts; aggregates predictions based on those models and determines the outcome based on the majority input.

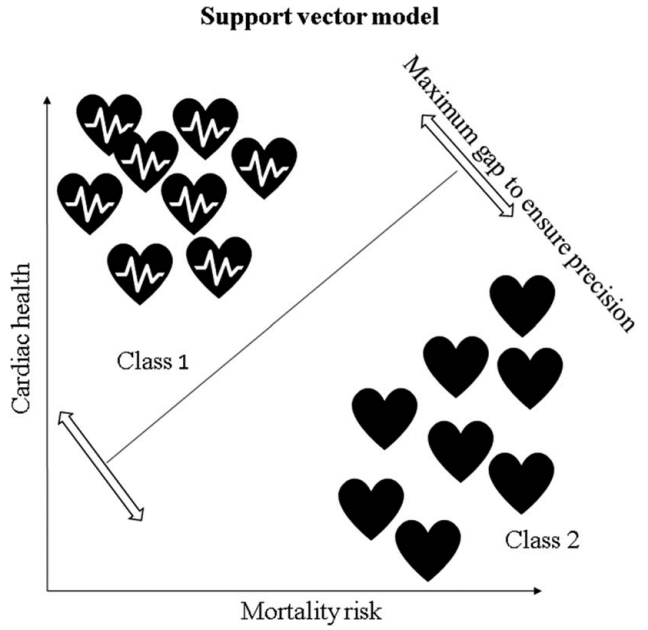

Figure 4. Patients separated into Class 1 (lower mortality risk) and Class 2 (higher mortality risk) using a support vector model. This line is strategically positioned to maximize the separation between the two groups, ensuring precise classification.

supervised ML algorithms. The performance of different algorithms was compared, revealing that support vector machines (SVM) exhibited desirable performance, achieving an area under the receiver operating characteristic (AUROC) curve of 0.77 and high accuracy during training and testing. ML algorithms and conventional statistical LR models identified similar predictors of all-cause mortality in HF, as revealed by the study[31]. The SVM is a valuable tool for predicting outcomes in HF patients. It employs a technique of drawing a line to separate patients who may encounter a particular outcome, such as mortality, from those who are less likely to. This line is strategically positioned to maximize the separation between the two groups, ensuring precise classification that is schematically presented in Figure 4. By utilizing SVM, healthcare professionals can identify patients with a higher risk of adverse outcomes, facilitating informed treatment decisions. This approach enables the distinction between HF patients with a higher likelihood of mortality and those with a lower likelihood, enabling the implementation of personalized patient management strategies. Both SVM and LR models identified various clinical factors as predictors, including medications (such as furosemide, beta-blockers, and spironolactone), physical examination findings (such as early diastolic murmur and parasternal heave), and comorbidities (such as coronary artery disease and ischemic cardiomyopathy). These predictors align with previous studies that have reported similar associations between clinical parameters and mortality in HF[31]. The study emphasized the significance of data quality and outlined plans to enhance the sample size by collecting data from multiple cardiac centers in sub-Saharan Africa, aiming to improve model performance[31].

Mortality in the ICU

In a retrospective cohort study conducted by Li *et al.*[29], ML algorithms were developed and validated for predicting the mortality of patients with HF in an ICU setting. Four ML algorithms, namely XGBoost, LR, random forest (RF), and SVM, were developed and compared. Superior performance was observed in the XGBoost model compared to the other algorithms. The XGBoost model was further explained using the SHAP (SHapley Additive exPlanations) method, which enhanced clinical interpretability and provided insights into the decision-making process. The study focused on a threshold probability range of 15–25% to prevent ineffective clinical interventions[32]. This range helped guide decision-making based on predicted probabilities. Setting a threshold probability range is akin to establishing boundaries to determine the likelihood of an event. For example, if the predicted probability for a person is above 0.7 within a threshold range of 0.3–0.7, it would be considered likely that they have the disease. On the other hand, if the predicted probability falls below 0.3, it would be deemed unlikely. Predicted probabilities between 0.3 and 0.7 would be regarded as uncertain and would require further investigation. This approach simplifies decision-making and categorizes individuals into different groups based on their predicted probabilities, enabling informed actions or further examination. Within the specified threshold range, the XGBoost model outperformed the other strategies examined. The most significant predictor variable was found to be the average blood urea nitrogen (BUN) level. Previous studies have also indicated the importance of BUN as a key predictor of HF mortality when using ML algorithms. BUN serves as a marker of renal function and reflects neurohormonal activation in HF patients. However, the study highlights the need for further research to explore the applicability of the SHAP method, as external validation was lacking.

In another study, an interpretable ML-based risk stratification tool for in-hospital all-cause mortality in ICU patients with HF was developed and validated by Chen *et al.*[31]. The authors compared their ML model with traditional risk prediction methods and highlighted the clinical implications of accurate prognostic evaluation in ICU patients with HF. ML enables the capture of both linear and nonlinear relationships from high-dimensional datasets compared to traditional generalized linear regression models, which fail to capture complex relationships between risk prediction factors and mortality endpoints in HF.

Analyzing and modeling such datasets require tailored approaches to effectively extract meaningful information and overcome dimensionality challenges. Chen *et al.*[31] utilized the XGBoost algorithm to develop a novel model that considered clinical features, comorbidities, and medication information. The model exhibited improved prediction performance compared to the Logistic model and the Get With The Guidelines-Heart Failure (GWTG-HF) model for all-cause mortality risk in ICU patients with HF. Through LASSO screening, the authors selected 17 easily accessible variables with high predictive value, facilitating risk assessment in primary hospitals. The SHAP value was employed to optimize prediction and interpretability of the XGBoost model, providing visual interpretation of the model's decision-making process. Furthermore, a website calculator was developed by the authors to assist physicians in understanding key features and prediction results[33].

Mortality in heart failure with preserved ejection fraction (HFpEF)

Gao *et al.*[32] conducted a comprehensive study to explore the prognostic value of circulating biomarkers in patients hospitalized for HFpEF. HFpEF is a common type of HF characterized by a preserved ejection fraction and diastolic dysfunction. The authors aimed to identify biomarkers that could predict the risk of all-cause death and cardiovascular death in these patients. Using Cox proportional hazards models, the researchers examined various circulating biomarkers to assess their association with the 2-year risk of death in HFpEF patients. They found several biomarkers to be significantly correlated with long-term mortality risk. One noteworthy finding was the independent predictive value of endoglin, a membrane co-receptor for transforming growth factor-b. Elevated levels of endoglin in the circulation were associated with inflammation, endothelial dysfunction, cardiac fibrosis, and vascular remodeling. These findings suggest that endoglin can serve as an important prognostic marker for HFpEF patients, reflecting the severity of cardiac impairment and predicting long-term mortality risk[34]. To further improve risk prediction in HFpEF, the authors developed a prediction model based on ML techniques[34]. The model was implemented using the SVM method, a powerful algorithm for classification and regression analysis. The results demonstrated the model's ability to accurately predict the 2-year risk of all-cause death in patients with acute HFpEF. This suggests that incorporating multibiomarker models based on ML can enhance risk stratification and provide valuable insights for clinicians managing HFpEF patients. The practical application of the multibiomarker prediction model holds promising potential in clinical practice. Analytical platforms capable of quantifying multiple protein biomarkers simultaneously using small plasma samples are already available, making multibiomarker tests affordable and accessible to most patients. Implementing this model in routine clinical practice could improve risk assessment and aid in personalized treatment strategies for HFpEF patients. Gao *et al.*'s[32] study sheds light on the prognostic value of circulating biomarkers in HFpEF patients. The development of a ML-based prediction model further enhances risk stratification, offering potential benefits for clinical decision-making in HFpEF management. Continued research in this field and the translation of these findings into clinical practice hold promise for improving outcomes in HFpEF patients.

In another study, Zhou *et al.*[35] explored the use of ML methods to predict the survival status of patients with HFpEF based on gene expression data. They focused specifically on HFpEF, which distinguishes their study from previous research that predominantly focused on predicting outcomes in HF in general. They compared six different prediction models and found that the Genetic Algorithm-Kernel Partial Least Squares (GA-KPLS) model, utilizing gene expression data, showed high accuracy in predicting survival status in HFpEF patients[35]. To illustrate the GA-KPLS model, let's consider a HF cardiologist using it to predict hospital readmission in HF patients just as an example. The model analyzes various patient factors, such as age, blood pressure, kidney function, and medication usage. It identifies age and kidney function as the most influential factors. The KPLS algorithm then examines the relationship between these factors and hospital readmission, revealing that older age and impaired kidney function are strongly associated with a higher likelihood of hospital readmission in HF patients as described in Figure 5. By applying the GA-KPLS model, the cardiologist can predict which HF patients are at a greater risk of being readmitted to the hospital based on their age and kidney function. This knowledge allows for proactive measures such as closer monitoring, medication adjustments, or specialized interventions to be implemented for these high-risk patients, ultimately reducing hospital readmissions and improving patient outcomes. The GA-KPLS model captures nonlinear relationships using kernel functions, resulting in more accurate predictions. This emphasizes the advantage of ML techniques in capturing complex patterns within genomic data. The potential applications of risk prediction models based on the GA-KPLS model are also discussed, including motivating patients to adhere to treatments, assisting clinicians in treatment decisions for high-risk patients, and informing the design of future HFpEF clinical trials[36].

Zhao *et al.*[37] conducted a study to develop and validate a risk prediction model, called HFmeRisk, for early assessment of HFpEF using data from the Framingham Heart Study (FHS) cohort. HFmeRisk combines different methods, including epigenetic factors and environmental exposures, to provide valuable insights for risk assessment of HFpEF at an early stage[37]. To illustrate two of the methods used in HFmeRisk, let's consider a example. Suppose we have a dataset with features such as age, blood pressure, cholesterol levels, and genetic markers for a group of patients. The LASSO algorithm is employed to identify the most relevant features by shrinking the less important ones towards zero. It helps determine which factors, such as age and blood pressure, have a stronger association with the risk of HFpEF while downplaying less significant factors. On the other hand, the XGBoost algorithm combines multiple weak models (decision trees) to create a more powerful predictive model. It iteratively learns from the data, giving more importance to samples that are difficult to predict correctly. Zhao *et al.*[37] reported how XGBoost could discover complex relationships between features and the risk of HFpEF, such as the combined impact of high cholesterol levels and specific genetic markers. By incorporating these methods into the HFmeRisk model, along with the deepFM algorithm that learns from data patterns, the tool offered a comprehensive assessment of early risk for HFpEF. It equips clinicians with valuable information to make informed decisions and take proactive measures in preventing or managing HF in patients.

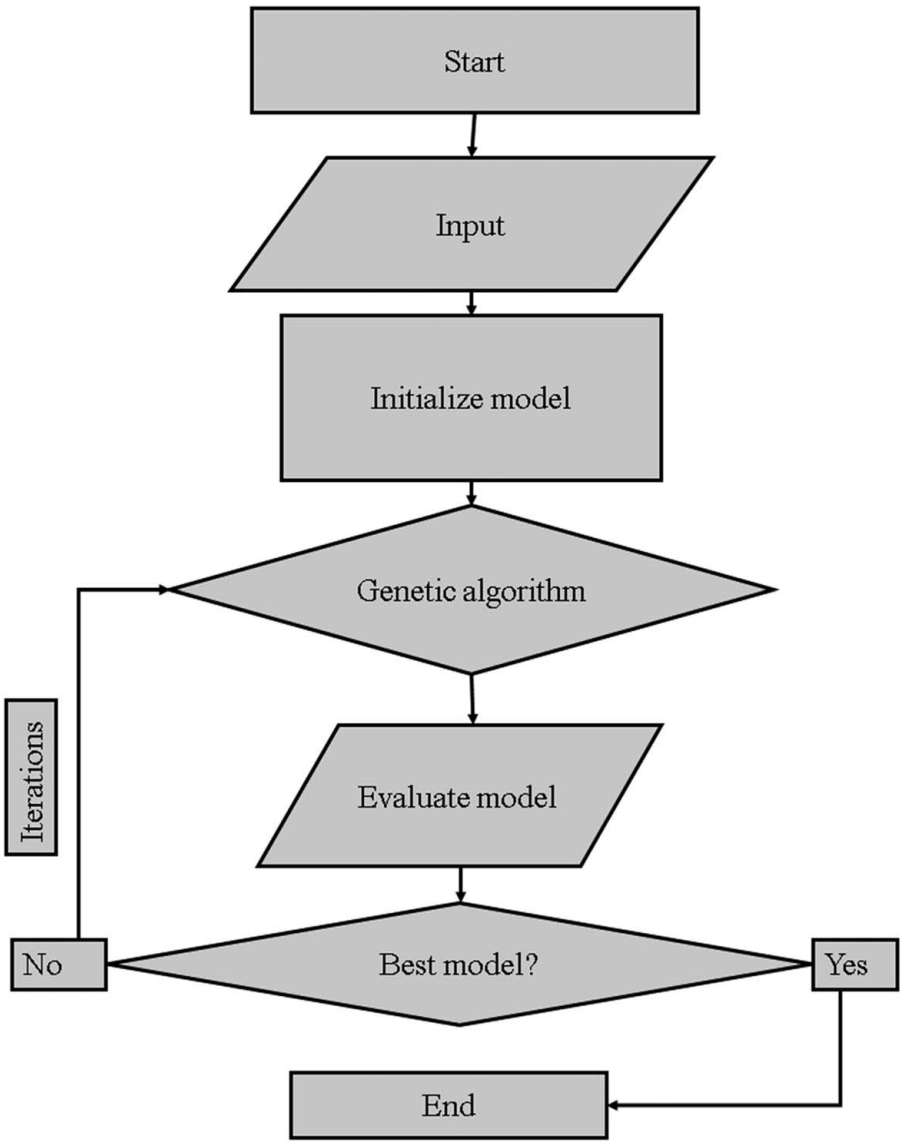

Figure 5. This flowchart illustrates the steps of the Genetic Algorithm-Kernel Partial Least Squares (GA-KPLS) algorithm. The algorithm starts with inputting the data and initializing the model, followed by the application of a genetic algorithm to generate variations of the model. Each variation is evaluated based on its fit to the data, and the best model is updated accordingly. This iteration continues until the best model is found. The flowchart showcases the iterative process of the GA-KPLS algorithm, demonstrating how it refines the initial model and helps discover meaningful relationships between variables in the data.

Mortality and hospitalization in patients with heart failure with mildly reduced ejection fraction (HFmrEF)

In their study, Zhao *et al*.[36] focused on developing and validating risk prediction models for HF hospitalization and all-cause mortality in patients with mildly reduced ejection fraction (HFmrEF). This subtype of HF is characterized by an ejection fraction ranging from 41 to 49%. Zhao *et al*.[36] developed eight alternative risk models. They employed ML techniques such as RF and LASSO regression to construct these models. The models incorporated various easily measurable clinical risk factors, making them practical for use in clinical practice. The study revealed that the physical condition of patients, as assessed by the Kansas City Cardiomyopathy Questionnaire (KCCQ) scores, was a strong predictor of both mortality and HF readmission over a 6-year follow-up period. Combining the KCCQ scores with NT-proBNP, a biomarker for HF, provided a quick risk assessment tool for HFmrEF patients. The study underscores the significance of utilizing ML techniques to improve the accuracy of risk prediction and uncover novel relationships between risk factors and outcomes. The developed models demonstrated good predictive performance for mortality and readmission in HFmrEF patients[36].

Readmission in patients with heart failure

In a study by Shin *et al*.[38], the predictive performance of ML methods with conventional statistical models (CSMs) in predicting readmission and mortality was compared among HF patients. The findings indicated that ML methods generally outperformed CSMs in these predictions, with tree-type ML algorithms being the most commonly used approach, and LR being the most frequent CSM approach. The practical applications of ML in HF prognostication

hold promise. ML offers advantages such as flexibility and nonparametric modeling. Nonparametric modeling allows for greater flexibility in capturing complex relationships and patterns without relying on predetermined mathematical functions or assumptions about the underlying data distribution. To illustrate the concept, consider a HF study where various patient characteristics are used to predict the risk of HF. In parametric modeling, specific mathematical equations, such as linear regression, are assumed to describe the relationship between these features and the risk of HF. However, nonparametric modeling takes a different approach by allowing the data itself to guide the analysis, adapting to data patterns using techniques like decision trees or RFs. Nonparametric models, like decision trees, can capture complex and nonlinear relationships between patient characteristics and the risk of HF. They adapt to intricate relationships that may not be easily described by predetermined equations. This adaptability is particularly useful when the true underlying data distribution is not well-defined or when relationships are intricate and difficult to specify in advance[38].

Sharma *et al.*[39] explored the use of administrative health data and ML models to predict the risk of unplanned readmissions in HF patients within 30 days after discharge. They compared the performance of ML models with the commonly used LaCE score, which includes four predictors. The ML models, leveraging more data and predictors, demonstrated improved predictive capabilities compared to the LaCE score. However, the study revealed that predicting readmissions in HF patients, whether using ML or non-ML methods, remains challenging. While the ML models outperformed the LaCE score, even the best-performing ML model provided only weak to moderate informative value as a classifier. They reported that ML models incorporating feature importance and impact plots, such as SHAP plots, can offer interpretability by identifying variables strongly associated with readmission risk[39].

Major adverse cardiovascular events (MACE) in patients with heart failure

Sun *et al.*[40] developed a scoring system to predict the risk of major adverse cardiovascular events in patients with congestive heart failure (CHF). Sun *et al.*[40] study aimed to develop a predictive model suitable for different HF populations, focusing on LVEF analysis variables without categorizing LVEF subtypes. The study introduced the use of velocity flow mapping (VFM) parameters in the predictive model. VFM parameters analyze blood flow patterns within the heart and provide insights into the likelihood of developing HF. By incorporating these parameters into the model, the risk of HF can be estimated. These parameters are obtained through color Doppler echocardiography, enabling the visualization of blood flow as a velocity vector. The model is trained on extensive patient data and patterns associated with increased risk are identified. When VFM measurements are inputted, the model provides an assessment of the patient's likelihood of developing HF. The model demonstrated good performance and showed significant differences in survival curves among different risk groups. ML algorithms, particularly the XGBoost classifier, were employed to analyze the predictive ability of the model. The inclusion of general parameters, speckle-tracking echocardiography (STE)-related parameters, and VFM-related parameters in the XGBoost classifier resulted in improved classification accuracy compared to other algorithms[40].

Heart failure prognosis

HF has a bad prognosis, according to studies of people hospitalized to hospitals with HF and treatment trials. This knowledge, however, only applies to a specific group of HF patients. For example, it is evident those clinical studies in HF have primarily involved men and that patients in these trials are younger and have fewer comorbidities than the normal community HF patient. There is little information available in society as a whole on the prognosis of HF; three studies looked at the prognosis in people who had been diagnosed with HF (i.e. prevalent HF, but not always new instances of HF). Furthermore, despite our study's significant response rate (79%), nonresponse may have resulted in an inaccurate estimation of HF survival, as the rates of responses were lower in older age groups and it is possible that those with severe HF were more unlikely to participate[41].

Tian *et al.*[42] conducted a study to evaluate prognostic models for patients with CHF using patient-reported outcomes (PROs). They developed and validated models that utilize PRO data to predict events like mortality and HF readmission in CHF patients. These models showed promising performance and can be easily implemented in clinical practice since they only require variables that can be collected after discharge. Tian *et al.*[42] emphasized the predictive value of PROs, the potential of ML methods with parameter adjustment, and the superiority of the XGBoost algorithm in their study. They also introduce interpretability techniques and a web-based risk calculator to enhance understanding and facilitate clinical decision-making. These findings contribute to the growing body of evidence supporting the integration of PROs into prognostic models for CHF patients, ultimately improving patient care and outcomes[42].

Gandin *et al.*[43] developed and compared two prognostic models for HF in diabetic patients using electronic health records (EHRs) and compared their performance to the Risk Equations for Complications Of type 2 Diabetes (RECODe). The first model employed a Cox proportional hazards model with elastic net regularization, while the second model utilized a deep neural network (DNN). Both models demonstrated superior performance compared to the RECODe risk equations. The DNN model exhibited moderate performance in terms of discrimination and well-calibrated predictions. It selected eight covariates that were either predictors of the RECODe equations or established risk predictors from previous studies. Interestingly, the DNN model indirectly incorporated atrial fibrillation through related variables derived from ECG features associated with atrial fibrillation. The DNN model exhibited adequate calibration, which is crucial for effective clinical decision-making. This study highlights the potential of leveraging EHRs and AI techniques to develop accurate prognostic models for HF in diabetic patients. The findings contribute to the growing body of knowledge on predictive modeling in healthcare and pave the way for further advancements in personalized risk assessment and management strategies for patients with diabetes and HF[43].

The various applications of AI are summarized in Table 1.

Research in heart failure

Traditional clinical trials often lack generalizability due to strict eligibility criteria, specialized environments, and limited data on real-world interactions and adherence to therapy. To offer a potential solution to this, D'Amario *et al.*[44] described the GENERATOR HF DataMart, an AI laboratory that generates real-world evidence for HF patients using real-world data (RWD). They highlighted the need for big data analytics and AI to

Table 1

This table summarizes various studies showcasing the application of AI in HF diagnosis, prediction, and prognosis.

Application	Study
Heart failure diagnosis	Unterhuber *et al.*[26] developed a CNN model for distinguishing between HFpEF and controls, achieving high discriminatory ability
	Wang *et al.*[27] utilized DL with LSTM to identify high-risk HFrEF patients using a large US nationwide commercial insurance dataset
	Zhou *et al.*[35] used ML methods to predict survival status in HFpEF patients based on gene expression data, demonstrating high accuracy
	Gao *et al.*[32] explored the prognostic value of circulating biomarkers in hospitalized HFpEF patients, developing a SVM-based prediction model
	Zhao *et al.*[36] developed and validated risk prediction models for HFmrEF patients, incorporating easily measurable clinical risk factors
Heart failure prediction	Tohyama *et al.*[28] emphasized ML's effectiveness in predicting prognosis for HF patients, developing a new prediction model called SMART-HF
	Mpanya *et al.*[30] trained six supervised ML algorithms to predict all-cause mortality in HF patients, with SVM exhibiting desirable performance
	Li *et al.*[29] developed ML algorithms for predicting mortality of HF patients in an ICU setting, with XGBoost demonstrating superior performance
	Chen *et al.*[31] developed an interpretable ML-based risk stratification tool for in-hospital mortality in ICU patients with HF, outperforming traditional methods
Heart failure prognosis	Tian *et al.*[42] developed and validated prognostic models for CHF patients using PROs and ML methods, showing promising performance and easy implementation
	Gandin *et al.*[43] developed prognostic models for HF in diabetic patients using EHRs and DNN, demonstrating superior performance compared to traditional methods
	Zhao *et al.*[36] evaluated prognostic models for HF patients using PROs and ML methods, emphasizing the predictive value of PROs and the superiority of XGBoost

CNN, convolutional neural network; DL, deep learning; DNN, deep neural network; EHR, electronic health record; HF, heart failure; HFmrEF, heart failure with mildly reduced ejection fraction; HFpEF, heart failure with preserved ejection fraction; HFrEF, heart failure with reduced ejection fraction; LSTM, long short-term memory; ML, machine learning; PROs, patient-reported outcomes; SMART-HF, Self-care Management Intervention in Heart Failure; SVM, support vector model.

manage the volume, velocity, variety, veracity, and value of cardiovascular data. The DataMart leverages RWD collected during routine clinical practice, providing a broader representation of real-life populations. The use of RWD and AI is an exciting area of research that has the potential to overcome the limitations of traditional clinical trials and improve the generalizability of study findings. However, careful consideration of data quality, accuracy, and reliability is necessary to ensure that the generated evidence is robust and reliable[44].

Conclusion

In conclusion, this review underscores the pivotal role of ML in transforming various aspects of HF management. From enhancing diagnostic accuracy, particularly in distinguishing between HF phenotypes like HFpEF and HFrEF, to facilitating risk prediction and prognosis, ML offers a promising avenue for personalized treatment strategies. By harnessing advanced computational techniques and analyzing vast clinical datasets, ML holds the potential to revolutionize HF care by providing actionable insights for improved patient outcomes. Future research should focus on refining ML algorithms, validating predictive models across diverse patient cohorts, and integrating ML-based approaches into routine clinical practice, ultimately leading to better outcomes and enhanced quality of life for individuals living with HF.

Ethical approval

Ethical approval was not required for this review.

Consent

Informed consent was not required for this review.

Sources of funding

The authors declare that there are no conflicts of interest.

Author contribution

M.S.: conceptualization; M.S., H.M., and P.P.: methodology, writing – original draft; M.S., A.M., and H.M.: visualization; M.S., P.P., H.M., A.M., F.N., S.A., and S.T.: writing – review and editing. All authors fulfill the ICMJE criteria for authorship. All authors reviewed the final version of the manuscript and approved it for submission and publication.

Conflicts of interest disclosure

No conflicts of interest.

Research registration unique identifying number (UIN)

This review has been registered in Open Science framework with the unique code osf.io/ymvsp Available online from: https://doi.org/10.17605/OSF.IO/YMVSP.

Guarantor

Hassan Mumtaz.

Data availability statement

Available upon reasonable request.

Provenance and peer review

Not commissioned, externally peer-reviewed.

Acknowledgements

None.

References

[1] Tripoliti EE, Papadopoulos TG, Karanasiou GS, *et al*. Heart failure: diagnosis, severity estimation and prediction of adverse events through machine learning techniques. Comput Struct Biotechnol J 2017;15: 26–47.

[2] Yancy CW, Jessup M, Bozkurt B, *et al*. 2017 ACC/AHA/HFSA focused update of the 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Task Force on clinical practice guidelines and the Heart Failure Society of America. Circulation 2017;136:e137–61.

[3] Yancy CW, Jessup M, Bozkurt B, *et al*. 2017 ACC/AHA/HFSA focused update of the 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. J Card Fail 2017;23:628–51.

[4] Metra M, Lucioli P. Corrigendum to 'Prevalence of heart failure and left ventricular dysfunction in China: the China Hypertension Survey, 2012-2015' [Eur J Heart Fail 2019;21:1329-1337. Eur J Heart Fail 2020;22:759.

[5] Roth GA, Johnson C, Abajobir A, *et al*. Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015. J Am Coll Cardiol 2017;70:1–25.

[6] Kim MS, Lee JH, Kim EJ, *et al*. Korean guidelines for diagnosis and management of chronic heart failure. Korean Circ J 2017;47:555–643.

[7] Ponikowski P, Voors AA, Anker SD, *et al*. 2016 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: the task force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur J Heart Fail 2016;18:891–975.

[8] Shah KS, Xu H, Matsouaka RA, *et al*. Heart failure with preserved, borderline, and reduced ejection fraction: 5-year outcomes. J Am Coll Cardiol 2017;70:2476–86.

[9] Angraal S, Mortazavi BJ, Gupta A, *et al*. Machine learning prediction of mortality and hospitalization in heart failure with preserved ejection fraction. JACC Heart Fail 2020;8:12–21.

[10] Kapoor M, Rossor AM, Laura M, *et al*. Clinical presentation, diagnosis and treatment of TTR amyloidosis. J Neuromuscul Dis 2019;6:189–99.

[11] Manolis AS, Manolis AA, Manolis TA, *et al*. Cardiac amyloidosis: an underdiagnosed/underappreciated disease. Eur J Intern Med 2019;67: 1–13.

[12] Alkhawam H, Patel D, Nguyen J, *et al*. Cardiac amyloidosis: pathogenesis, clinical context, diagnosis and management options. Acta Cardiol 2017;72:380–9.

[13] Kogut J, Popjes ED. Hypertrophic cardiomyopathy 2020. Curr Cardiol Rep 2020;22:154.

[14] Ryu AJ, Kumar V, Borlaug BA, *et al*. Systolic-to-diastolic myocardial volume ratio as a novel imaging marker of cardiomyopathy. Int J Cardiol 2021;322:272–7.

[15] Michaud M, Maurin V, Simon M, *et al*. Patients with high left ventricular filling pressure may be missed applying 2016 echo guidelines: a pilot study. Int J Cardiovasc Imaging 2019;35:2157–66.

[16] Richards AM. N-terminal B-type natriuretic peptide in heart failure. Heart Fail Clin 2018;14:27–39.

[17] Vaduganathan M, Claggett BL, Desai AS, *et al*. Prior heart failure hospitalization, clinical outcomes, and response to sacubitril/valsartan compared with valsartan in HFpEF. J Am Coll Cardiol 2020;75:245–54.

[18] Anjan VY, Loftus TM, Burke MA, *et al*. Prevalence, clinical phenotype, and outcomes associated with normal B-type natriuretic Peptide levels in heart failure with preserved ejection frac-tion. Am J Cardiol 2012;110:870–6.

[19] Pieske B, Tschope C, de Boer RA, *et al*. How to diagnose heart failure with preserved ejection fraction: the HFA-PEFF diagnostic algorithm: a consensus recommendation from the Heart Failure Association (HFA) of the European Society of Cardiology (ESC). Eur J Heart Fail 2020;22: 391–412.

[20] Reddy YN, Carter RE, Obokata M, *et al*. A simple, evidence-based approach to help guide diagnosis of heart failure with preserved ejection fraction. Circulation 2018;138:861–70.

[21] Xanthopoulos A, Giamouzis G, Skoularigis J, *et al*. Heart failure with reduced, mildly reduced, or preserved left ventricular ejection fraction: Has reasoning been lost? World J Cardiol 2022;14:438–45.

[22] Cook JA, Collins GS. The rise of big clinical databases. Br J Surg 2015; 102:e93–101.

[23] Averbuch T, Sullivan K, Sauer A, *et al*. Applications of artificial intelligence and machine learning in heart failure. Eur Heart J Digit Health 2022;3:311–22.

[24] Wang J. Heart failure prediction with machine learning: a comparative study. J Phys: Conf Ser 2021;2031:012068.

[25] Yasmin F, Shah SMI, Naeem A, *et al*. Artificial intelligence in the diagnosis and detection of heart failure: the past, present, and future. Rev Cardiovasc Med 2021;22:1095–113.

[26] Unterhuber M, Rommel KP, Kresoja KP, *et al*. Deep learning detects heart failure with preserved ejection fraction using a baseline electrocardiogram. Eur Heart J Digit health 2021;2:699–703.

[27] Wang Z, Chen X, Tan X, *et al*. Using deep learning to identify high-risk patients with heart failure with reduced ejection fraction. J Health Econ Outcomes Res 2021;8:6–13.

[28] Tohyama T, Ide T, Ikeda M, *et al*. Machine learning-based model for predicting 1 year mortality of hospitalized patients with heart failure. ESC Heart Failure 2021;8:4077–85.

[29] Li J, Liu S, Hu Y, *et al*. Predicting mortality in intensive care unit patients with heart failure using an interpretable machine learning model: retrospective cohort study. J Med Internet Res 2022;24:e38082.

[30] Mpanya D, Celik T, Klug E, *et al*. Predicting in-hospital all-cause mortality in heart failure using machine learning. Front Cardiovasc Med 2022;9:1032524.

[31] Chen Z, Li T, Guo S, *et al*. Machine learning-based in-hospital mortality risk prediction tool for intensive care unit patients with heart failure. Front Cardiovasc Med 2023;10:1119699.

[32] Gao Y, Bai X, Lu J, *et al*. Prognostic value of multiple circulating biomarkers for 2-year death in acute heart failure with preserved ejection fraction. Front Cardiovasc Med 2021;8:779282.

[33] Kapur NK, Wilson S, Yunis AA, *et al*. Reduced endoglin activity limits cardiac fibrosis and improves survival in heart failure. Circulation 2012; 125:2728–38.

[34] Kapur NK, Heffernan KS, Yunis AA, *et al*. Usefulness of soluble endoglin as a noninvasive measure of left ventricular filling pressure in heart failure. Am J Cardiol 2010;106:1770–6.

[35] Zhou L, Guo Z, Wang B, *et al*. Risk prediction in patients with heart failure with preserved ejection fraction using gene expression data and machine learning. Front Genet 2021;12:652315.

[36] Zhao H, Li P, Zhong G, *et al*. Machine learning models in heart failure with mildly reduced ejection fraction patients. Front Cardiovasc Med 2022;9:1042139.

[37] Zhao X, Sui Y, Ruan X, *et al*. A deep learning model for early risk prediction of heart failure with preserved ejection fraction by DNA methylation profiles combined with clinical features. Clin Epigenetics 2022;14:11.

[38] Shin S, Austin PC, Ross HJ, *et al*. Machine learning vs. conventional statistical models for predicting heart failure readmission and mortality. ESC Heart Fail 2021;8:106–15.

[39] Sharma V, Kulkarni V, McAlister F, *et al*. Predicting 30-day readmissions in patients with heart failure using administrative data: a machine learning approach. J Card Fail 2022;28:710–22.

[40] Sun Q, Jiang S, Wang X, *et al*. A prediction model for major adverse cardiovascular events in patients with heart failure based on high-throughput echocardiographic data. Front Cardiovasc Med 2022;9: 1022658.

[41] Nikolov A, Popovski N. Extracellular matrix in heart disease: Focus on circulating collagen type I and III derived peptides as biomarkers of myocardial fibrosis and their potential in the prognosis of heart failure: a concise review. Metabolites 2022;12:297.

[42] Tian J, Yan J, Han G, *et al*. Machine learning prognosis model based on patient-reported outcomes for chronic heart failure patients after discharge. Health Qual Life Outcomes 2023;21:31.

[43] Gandin I, Saccani S, Coser A, *et al*. Deep-learning-based prognostic modeling for incident heart failure in patients with diabetes using electronic health records: a retrospective cohort study. PLoS ONE 2023;18: e0281878.

[44] D'Amario D, Laborante R, Delvinioti A, *et al*. Generator heart failure datamart: an integrated framework for heart failure research. Front Cardiovasc Med 2023;10:1104699.