OXFORD

Precision Clinical Medicine, 2025, **8**: pbaf016
DOI: 10.1093/pcmedi/pbaf016
Review

Utilizing multimodal artificial intelligence to advance cardiovascular diseases

Xin-yue Yang [1,2,3,‡], Yi-ming Li,[1,2,3,‡] Jian-yong Wang,[4] Yu-heng Jia [5], Zhang Yi,[4,*] Mao Chen [1,2,3,*]

[1]Department of Cardiology, West China Hospital, Sichuan University, Chengdu 610041, China
[2]Laboratory of Cardiac Structure and Function, Institute of Cardiovascular Diseases, West China Hospital, Sichuan University, Chengdu 610041, China
[3]Cardiac Structure and Function Research Key Laboratory of Sichuan Province, West China Hospital, Sichuan University, Chengdu 610041, China
[4]Machine Intelligence Laboratory, College of Computer Science, Sichuan University, Chengdu 610040, China
[5]The Heart Center, Rigshospitalet, University of Copenhagen, Copenhagen 2100, Denmark
*Corresponding authors: Mao Chen, hmaochen@vip.sina.com; Zhang Yi, zhangyi@scu.edu.cn
‡Xin-yue Yang and Yi-ming Li contributed equally to this work.

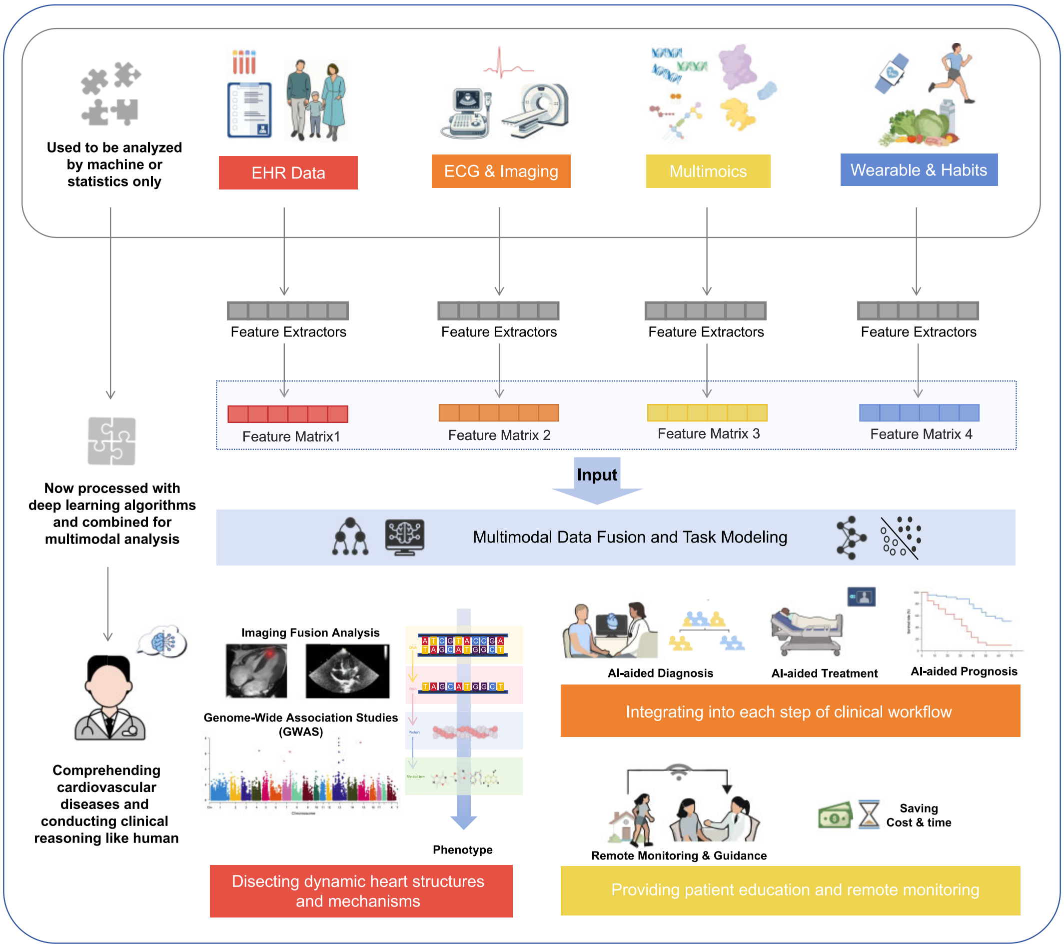
Abstract

The emergence of artificial intelligence (AI) is transforming cardiovascular medicine. Initially, AI applications concentrated on analyzing single data types, such as electrocardiograms and imaging studies. However, advancements in multimodal AI have now enabled the integration of diverse data sources, facilitating a comprehensive understanding of patient health and predictive accuracy of disease outcomes. In this review, we discuss current achievements in multimodal AI within cardiovascular medicine, including various combinations of different modalities, computer algorithms of data integration and fusion, and their integration into clinical workflow. As the field continues to evolve, we further propose current challenges and prospects for their future implementation.

Keywords: artificial intelligence; multimodal; cardiovascular medicine; cardiology

Received 17 May 2025; revised 6 July 2025; accepted 8 July 2025. published 17 July 2025

Graphical Abstract



Preprocessing, building, and applying multimodal artificial intelligence (AI) in cardiology.

Introduction

Cardiovascular disease (CVD), the leading cause of death globally, is a clinically complex condition whose diagnosis has long required multimodal data integration. From basic dissection and observation to understanding the heart's structures by William Harvey to simply hearing and interpreting heart sounds by René Laennec, now various data modalities—including laboratory tests, electrocardiograms, imaging techniques, and multiomics—can be assessed comprehensively. The diversity of medical data types and content enables deeper investigation into personalized diagnoses and effective treatments. Traditional linear models leverage several key clinical data including medical history, physical exam results, and select lab tests or imaging for risk prediction.

However, their inability to discover interconnections between different prognostic features or provide supplementary information impedes future management optimization. Moreover, quantitative analysis of high-dimensional features to inform clinical decision-making is time-consuming and labor-intensive for traditional tools.

This is where artificial intelligence (AI) comes in. Over the last decade, AI has demonstrated remarkable performance in numerous clinically relevant tasks [1]. This technology, encompassing data acquisition, preprocessing, feature extraction, and model construction, enhances our understanding of cardiac severity and heterogeneity. However, a main issue with AI applications is that most of them are unimodal, typically learning only specific patterns from training databases and missing comprehensive analysis of diseases from multiple perspectives. This potentially restricts their generalizability to interpret disease dynamic progression. The evolution of models like Transformer expands data volume and processing speed to fuel multimodal analysis, combining supplementary data and clinical context in a way more like real-world clinicians, generating more precise predictions and directing the identification of new biomarkers. Now, the field of cardiology is on the threshold of a new era in multimodal AI to incorporate into clinical practice [2, 3]. Multiple fusion of different modalities has been extensively witnessed, permitting multidimensional cardiac morphology quantification and pathophysiological analysis to advance clinical diagnostics and therapy. Our goal is to cover current research, emerging trends, and the future utilization of multimodal AI in the management of CVD in clinical practice (Fig. 1).

Why do we combine different modalities in CVD?

Diagnosing CVD requires multimodal approaches, yet each modality only excels in limited aspects. Echocardiography can reliably describe most structural abnormalities, but the echocardio-

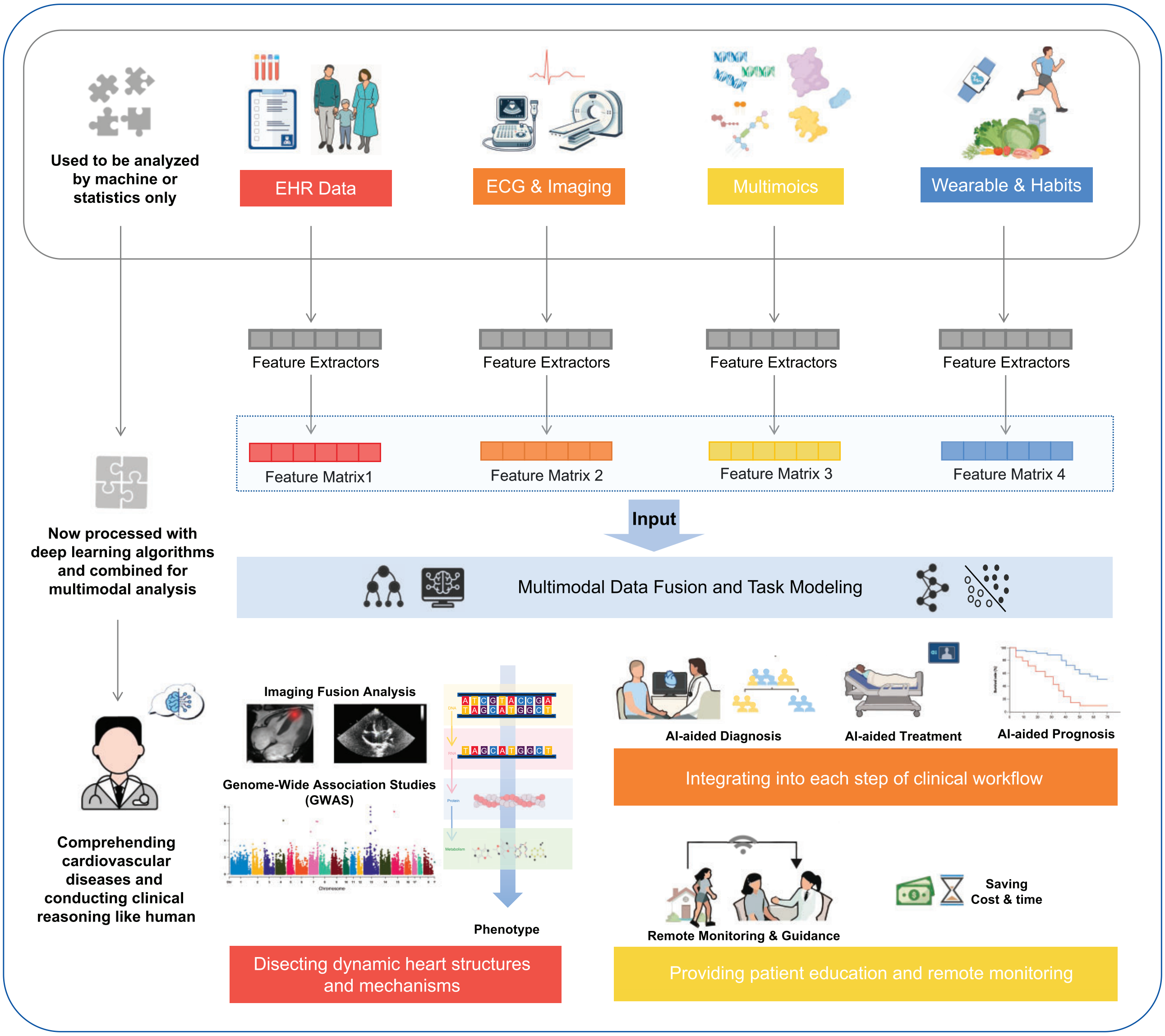


Figure 1. Preprocessing, building, and applying multimodal AI in cardiology. In conventional clinical practice, information collected from different modalities is usually separately analyzed by machine or statistics. With the rise of deep learning algorithms, multimodal data are now collected as input signals, and various methods are utilized in processing and selectively extracting features to build multimodal AI. Formed multimodal AI can be applied to different aspects of CVDs and stimulate clinical reasoning like a human. EHR, Electronic health record.

graphic assessment function is more challenging. Magnetic resonance imaging (MRI) excels in serial examinations of function, yet its clinical utility is constrained by prohibitive costs and prolonged scanning durations. Different combinations could complement each other to generate a more comprehensive analysis [4] (Table 1). For instance, the integration of invasive coronary angiography (ICA) with myocardial perfusion imaging (e.g. single photon emission computed tomography (SPECT) or positron emission tomography (PET)) enables refined hemodynamic assessment in coronary artery disease (CAD). This multimodal synergy allows detailed characterization of atherosclerotic plaques and the identification of those at risk of rupture could be recognized early [5]. To refine cardiac structural interpretation, echocardiography-fluoroscopic fusion imaging provides better understanding of the 3D relationship of anatomy and devices, improving the confidence of positioning and visual guidance during procedures [6]. However, protracted data processing constrains scalability, driving the need for computational acceleration.

How is the combination of different modalities achieved with multimodal AI?

Multimodal AI, a pivotal research domain within AI, has undergone progressive development. The rise of machine learning (ML) enables computer systems to learn from data and improves their performance over time without being explicitly programmed. Deep learning (DL), going further, features in neural networks with multiple layers to automatically deal with unstructured and high-dimensional data, showing great advantages with large datasets. AI, in a broader sense, covers various fields such as robotics, natural language processing, ML, DL, and expert systems. Multimodal data analysis is a pivotal research domain within AI, focusing on feature engineering, feature fusion, and decision-making, greatly enhancing physicians' ability to cope with vast amounts of information [7]. Modalities included can be of different types (such as images, text, and audio), or of the same type but with different characteristics or sources [such as computed tomography (CT)

Table 1. Examples of current multimodal data fusion in CVDs.

Echocardiography + CT/CCTA	Assess cardiac structure and function, providing morphological characteristics. Echocardiography often suffers from poor acoustics and depends on the operating angle and techniques. Computed tomography (CT) is a rapid way to provide heart anatomy with relatively high resolution, used for diagnosing cardiomyopathy, heart failure, valvular diseases, etc. [99–101].
Echocardiography + SPECT/PET	Assess cardiac structure and function, as well as contractile function and wall motion, providing states of inflammation in the heart and guiding clinical treatment. For example, PET radiotracer techniques can be used to identify, monitor, and target molecular structures involved in heart failure progression, along with basic function analysis from echocardiography [102, 103].
Echoardiography + CMR	Assess cardiac structure and function, identifying myocardial tissue characteristics (e.g. myocardial fibrosis and scars) which cause diastolic dysfunction, together with parameters for filling pressure [104, 105].
CT + SPETCT/PET	Assess coronary blood flow, impact on the myocardium, and exclude coronary stenosis. CT, particularly CCTA, clearly shows the degree of coronary artery stenosis and plaque characteristics, providing detailed anatomical information. PET, on the other hand, offers a detailed assessment of myocardial perfusion and viability, offering functional insights. For example, CT can assess coronary artery stenosis through CCTA, while PET can evaluate myocardial blood flow through myocardial perfusion imaging [106].
CCTA + CMR	MRI excels in soft tissue contrast and myocardial viability assessment, which is crucial for diagnosing myocardial ischemia and myocardial infarction. CT is superior in displaying the anatomical structure of the heart, coronary artery calcification, and pulmonary lesions. Together, they help doctors accurately assess the degree of coronary artery stenosis and changes in heart morphology [107–109].
SPETCT/PET + CMR	MRI offers high-resolution anatomical details and visualizations of soft tissues with excellent contrast. PET/SPET provides functional and molecular insights, including myocardial perfusion, metabolic activity, and the presence of specific molecular markers. This integration enhances diagnostic accuracy by cross-validating findings and detecting subtle changes that might be missed by one modality alone, such as myocardial viability, inflammation, and atherosclerotic plaques. Additionally, the combined approach reduces ionizing radiation exposure and increases patient comfort by potentially allowing for simultaneous or sequential imaging in a single session [110, 111].

and MRI], at least two of which are supposed to be leveraged to form the final multimodal AI [8].

Neural networks and large language models for unified feature space mapping and deep reasoning

In recent years, deep neural networks have become foundational to multimodal data analysis due to their powerful capabilities in feature extraction and reasoning [9]. Tremendous cross-modal representations provide complementary and corroborative information across modalities [10] for feature selections and comprehensive analysis. Primary architectures proposed in this area often include fully connected neural networks (FCNN), convolutional neural networks (CNN), recurrent neural networks (RNN), and so on. The end-to-end learning framework of neural networks transcends traditional approaches to enable a higher-level, more abstract analysis of multimodal information.

Additionally, advancements in large language models and vision models have catalyzed significant progress in multimodal large language models (MLLMs). These models enable natural interactions between vision and language modalities, achieving semantic alignment and fusion across modalities by employing instruction-based fine-tuning and other training techniques that enhance multimodal response capabilities [11]. In clinical diagnostics, for example, datasets containing CT, ultrasound, MRI, and pathological images, alongside radiology reports, ultrasound summaries, and electronic health records (EHRs), offer a robust foundation and ideal application scenarios for MLLMs. Models like pathology copilot [12], EchoClip [13], SkinGPT-4 [14], and BiomedParse [15] have demonstrated impressive capabilities in zero-shot learning, adaptability, prompt-based learning, and multi-turn reasoning (Table 2). Together, they reveal intricate and nonlinear relationships both within and across various modalities, delivering innovative solutions for multimodal data analysis within the medical domain.

Methods of data fusion with multimodal AI

Figuring out how to obtain individual and combined representations of diverse modalities in a manner that allows for their successful integration is a core issue in multimodal analysis. Common methods applied include early fusion, intermediate fusion, and late fusion [16]. Early fusion involves combining raw data from multiple sources, like merging genetic information with imaging data at the outset, to create a comprehensive dataset before feature extraction. Intermediate fusion, on the other hand, integrates features extracted from individual modalities, allowing for a more nuanced integration after initial processing. Late fusion delays integration until the decision-making stage, where predictions from separate models trained on different data types, such as a model trained on patient history and another on blood biomarkers, are combined to make a final diagnosis. More data fusion methods are being exploited nowadays [17], such as attention mechanism methods [15] and encoder–decoder methods [10]. There is no conclusive evidence that one fusion type is ultimately better than the others, as each type is heavily data and task specific. As technology advances, the number of multimodal medical data fusion methods is expected to grow [17] (Table 3, Figs. 2 and 3).

Powering multimodal AI to depict the heart from morphology and physiology

The intricate nature of CVD lies in its complex anatomy. Alterations in these structures can result in severe heart defects. The application of multimodal AI to cardiac imaging helps to efficiently optimize anatomical, morphological, and functional data

Table 2. Examples of multimodal large language models applied in clinical medicine.

Model name	Modalities involved	Applications
PathChat [12]	Vision: pathology images Text: pathology text	PathChat is trained with pathology image and caption pairs to align its image, and fine-tuned by curated instructions, which allow it to answer complex pathology-related questions like "What is the diagnosis?" or "Can you briefly describe the image?", more accurately than pathologists.
EchoCLIP [13]	Vision: cardiac ultrasound videos Text: expert interpretations	EchoCLIP performs well on a diverse range of benchmarks for cardiac image interpretation, like assessing cardiac function and identifying implanted intracardiac devices.
SkinGPT-4 [14]	Vision: skin disease images Text: clinical concepts and doctors' notes	With SkinGPT-4, users could upload their own skin photos for diagnosis, and the system could autonomously evaluate the images, identify the characteristics and categories of the skin conditions, perform in-depth analysis, and provide interactive treatment recommendations.
BiomedParse [15]	Vision: images from 45 biomedical image segmentation datasets, including MRI, CT, PET, ultrasound imaging, SPECT, optical coherence tomography, X-rays, histopathological sections, and cell microscopy images Text: semantic information from these 45 datasets and descriptive labels synthesized by GPT-4	BiomedParse outperformed existing methods on image segmentation across nine imaging modalities, with larger improvement on objects with irregular shapes, and could simultaneously segment and label all objects in an image, such as brain tumors, pulmonary nodules, immune cells, and cancer cells, providing diagnostic evidence for clinicians.

Table 3. Advantages and disadvantages of commonly used multimodal fusion methods [7, 9, 17, 112, 113].

Fusion method	Introduction	Advantages	Disadvantages
Early fusion	The original input data (such as patient records, imaging data, etc.) are concatenated, and together, treated as a unimodal input feature. This means that the DL architecture does not differentiate from which modality the features originate.	It is simple to design and can learn cross-modal data relationships from low-level features.	This approach may fail to identify relationships between modalities that only emerge at higher levels of abstraction and is also sensitive to different sampling rates of modalities.
Intermediate fusion	Feature vectors are learned and fused instead of the original multimodal data, which means that each modality can be learned through neural networks of the same type or different types of networks, and then learned features are treated as inputs to the fusion layers are features.	It is flexible in finding the right depth and sequence of fusing each modality, which may reflect more closely the true relationships between the modalities, since highly correlated modalities are fused earlier and other modalities later in the architecture.	It requires careful design of the fusion mechanism to effectively combine features from different modalities. A poorly designed fusion module may fail to capture the relationships between modalities.
Late fusion	Separate models are trained for each modality to extract features, and decisions are made by the fusion of decisions by sub-models.	It is quite independent since the errors made by each model are not perfectly correlated or affected by the missed models. Decisions are made taking advantage of the complementary information from each modality, which is especially promising for the fusion of heterogeneous modalities.	It is not able to learn interactions between features of different modalities.

to be combined. By fusing MRI and CT images, accurate whole-heart geometries with good anatomical consistency and high resolution can be yielded [18]. Even with relatively low input qualities, they could still provide detailed information [19]. This kind of precise and robust structural identification lays a foundation for accurately tracking cardiac motion and turning the static model into live-moving hearts. By combining traditional radiology methods with the "time" factor, 2D to 3D images are successfully turned into 4D heart models, which can produce geometries that better satisfy modeling requirements for cardiac flow simulations and image-based strain measurements, paving the way for comprehensive assessment of cardiac function and morphology [20–22]. Electrophysiology of the heart can also be interpreted by DL signals from 12-lead ECG. Together with imaging methods, such as X-ray, cardiac accessory pathway, which may cause paroxysmal palpitations and occasionally fatal arrhythmias, can be exposed [23]. They may enable doctors to view, measure, and manipulate the patient's anatomical structures in real time, with imaging explicitly highlighting the specific structural changes leading to functional impairments (Fig. 4).

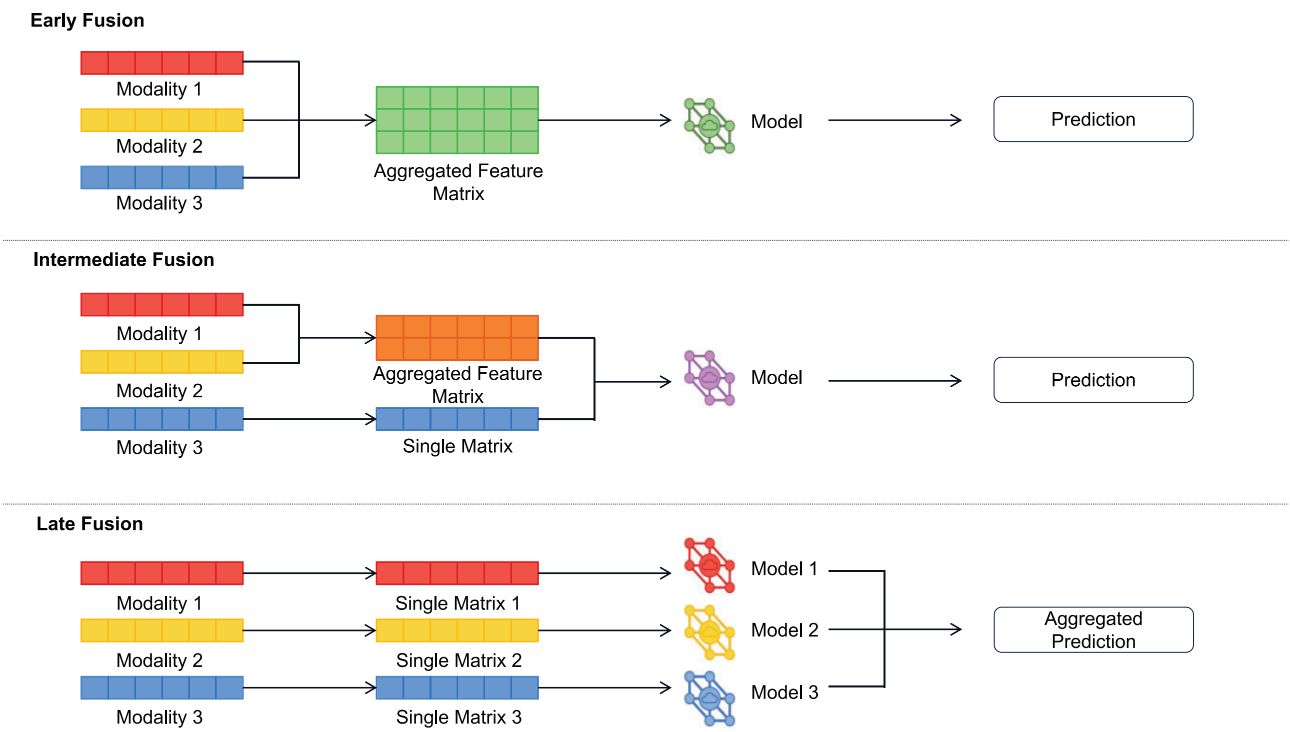


Figure 2. Three commonly used multimodality fusion frameworks. Early fusion combines information from all modalities at the input level for processing by a single model. Intermediate fusion combines modalities across representation layers simultaneously or progressively, allowing fusion of highly correlated modalities first before incorporating less correlated data in deeper layers. Late fusion trains separate models per modality and aggregates their predictions, making it ideal for heterogeneous data.

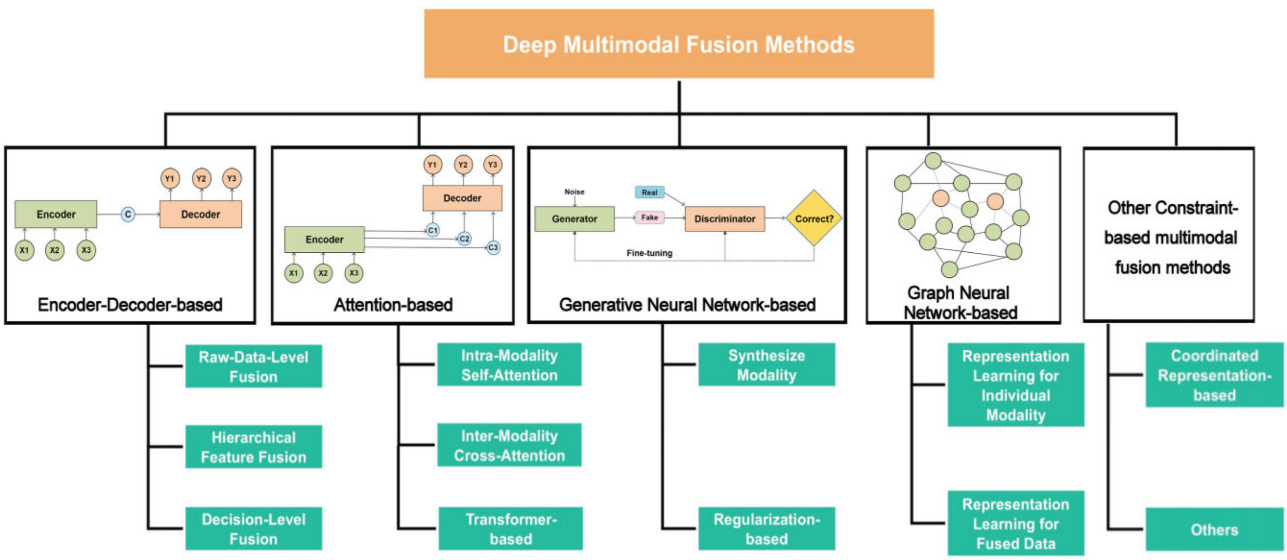


Figure 3. Deep neural network-based methods for multimodal fusion.

Applying multimodal AI analysis to reveal complex mechanisms

Other factors limiting our understanding CVD are the molecular mechanisms that direct cardiovascular homeostasis, cellular composition, and (micro-)environment, exacerbated by the complex structure. Indeed, the integration of DL imaging and systems biology is essential to complement comprehensive data-driven insights, linking radiomics phenotypes to gene phenotypes. Combining the U-net algorithm with available cardiac MRI and genotyping data from the UK Biobank, a genetic landscape that influences aortic forward velocity, left ventricular stroke volume, and aortic regurgitation fraction is unveiled [24]. When applying a similar method to defining quantitative traits of enlargement or aneurysm of the aorta, significant genetic variants were found associated with descending aortic diameter, aortic root size, and aortic valve regurgitation fraction [25]. Extending beyond the genomic level, multimodal AI is also extensively related to proteomics editing and activities of epigenetic regulation. Combined with genome-wide association analyses and phenome-wide association analyses, studies have started to embark on a profound

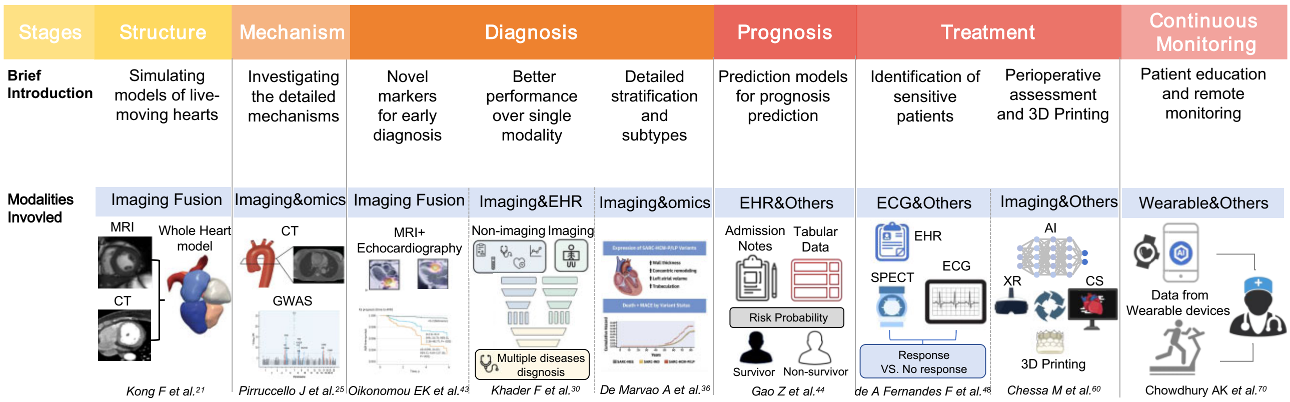


Figure 4. Schematic diagram of the application of multimodal AI in CVDs. It illustrates the multifaceted application of AI in the management of CVDs, including structural modeling, mechanism analysis, diagnosis, prognosis assessment, treatment strategies, and continuous monitoring. Representative examples and commonly used combinations of modalities are exhibited, with medical imaging fusion remaining the most extensively deployed. They could complement each other to form more comprehensive illustrations than human or single-modal data. GWAS, genome-wide association studies; XR, extended reality; CS, computational simulations. (Figure created and adapted from references listed above).

exploration of the complex interplay between genetic and phenotypic factors, providing a more detailed understanding of CVD mechanisms and family screening [26]. Key mechanisms involving pathways implicating metabolic dysregulation, inflammation, and fibrosis have also been revealed [27]. Together, they highlight a significant multiomics component in heart function, underpinning key unanswered aspects of microscopic pathogenesis and macroscopic structural changes, deepening our understanding of the mechanisms behind structure and function (Fig. 4).

Integrating advanced multimodal AI into clinical workflows

Improving diagnostic accuracy and subtype classification by breaking the monomodal barriers

With the emergence of multimodal AI, diagnostic precision based on new fusion of different data resources has been seen to outperform human readers [28]. These models prove that, after assessing time, frequency, energy, or entropy spent, optimal subsets of features can be identified to depict the disease [29], diagnosing multiple diseases at the same time. A transformer-based neural network to integrate multimodal patient data from chest radiographs and clinical parameters enhances diagnostic performance for up to 25 pathologies [30], achieving an average area under the curve (AUC) of 0.77 in diagnosing diseases within an intensive care unit (ICU) setting. Together, they highlight the advancements in multimodal cardiovascular AI as a powerful tool for disease diagnosis across various facets.

Besides, as different types are associated with different disease progression, subtyping is also crucial to diagnosis. Computational disease subtyping techniques can harness the power of data science and ML, unveiling patient groups within a high-dimensional space [31]. By meticulously integrating 645 distinct clinical factors with genetic information [32], 5 incident heart failure (HF) phenotypes, including early onset, late onset, atrial fibrillation-related, metabolic, and cardiometabolic, were identified, providing implications for future research and clinical practice. Deeper investigations are able to pinpoint the causal pathways specific to phenotypically distinct subpopulations [33]. Leveraging residual CNNs (ResNet200), another group delineated abdominal aortic aneurysm (AAA)-associated lesion regions on computed tomography angiography (CTA), revealing two subgroups [34]. Differences in the prognosis are in significant correlation with immune cell activation and AAA pathology, offering a novel therapeutic strategy for future management. Similar methods have also emerged in the study of hypertension [35]. To delve deeper into genetic relationships, a study utilized whole exome sequencing and MRI from participants within the UK Biobank [36], revealing three distinct stratifications based on their genetic profiles. This detailed classification allows for a nuanced understanding of the variable expressivity and penetrance of hypertrophic cardiomyopathy (HCM)-associated genes and their significant association with adverse cardiovascular outcomes, enhancing the provision of personalized medicine approaches in cardiomyopathy healthcare. As molecular data linked with rich clinical data are becoming easily accessible, multimodal AI for subtyping has the potential to advance our knowledge of how diseases are defined and treated (Fig. 4).

Identifying novel markers for early risk stratification and better prognosis

Multimodal biomedical AI, built from transformers, has shown considerable potential in enhancing early risk stratification and propels advancements in diagnostic technologies that offer cheap, rapid, and sensitive methods to improve the management of CVD. Current risk assessments have shortcomings. Although clinical risk factors are incorporated into tools like QRISK [37], SCORE [38], and China-PAR [39], these calculators often misestimate risk—especially in asymptomatic patients whose low risk factor scores do not meet diagnostic criteria [40].

To refine risk assessment, there is a need for a multi-biomarker approach for precision medicine. Multimodal AI can reveal multifactorial risk contributors, providing new grading systems for patients. Early on admission, description notes combined with tabular data can provide a relatively precise prognostic evaluation of hospital outcomes [41]. Physiological data from EHRs generates new markers as well, even extending into the pre-clinical stage. Another *in silico* quantitative marker for CAD, utilizing clinical features and recorded medications [42], is able to depict subclinical atherosclerosis to advanced disease states. Ranging from 0 (lowest probability) to 1 (highest probability), it can capture the continuum of CAD risk, non-invasively quantifying atherosclerotic plaque burden and risk of all-cause mortality. Similarly, using echocardiography or CMR imaging, a novel AI-based video biomarker was found to be independently associated with the development and progression of aortic stenosis (AS), facilitating opportunistic risk stratification [43].

As for prognosis and outcomes, admission notes and clinical tabular data are structured as a potentially novel method in evaluating the risk of mortality in patients with HF [44], reaching an AUC score of 0.767, providing more accurate and timely decision support than single modality data. Together, they show the ability of multimodal AI to identify novel biomarkers, successfully mapping the progression of the disease from pre-clinical stages to overt manifestation, driving data-driven predictions of the risk of a disease-related endpoint (Fig. 4).

Advancing multimodality-AI models to aid therapy responses and sensitive populations

More than static clinical endpoint, the multimodal AI method could also be utilized to identify potential populations that may benefit from the treatment, and thereby provide guidance for the selection of clinical treatments [45]. A model utilizing a comprehensive dataset identifies patients with secondary mitral regurgitation (SMR) suitable for transcatheter edge-to-edge repair (TEER) [46]. The identification of an extreme-risk population highlights the potential of models to guide clinical decision-making regarding the utility of TEER, identifying cases where the procedure may not be beneficial. This is especially important in cardiac resynchronization therapy (CRT) therapy as well. Despite clear guidelines for which patients should be treated and the precise process of treatment, a significant number of patients do not respond well, highlighting the need for better patient selection and treatment optimization. Through ML, AI coping with multimodal data may predict CRT response [47] and its relation to left ventricular ejection fraction (LVEF). In fact "responders" (increase in LVEF $\geq$ 5%) and "super-responders" (increase in LVEF $\geq$ 15%) with distinct features [48] have been revealed, providing novel insights into the clinical decision of CRT. Evaluation of drug responses nowadays, which is also predominantly derived from population-based evidence, can also benefit. The emergence of multimodal AI presents opportunities to identify specific patient subgroups likely to benefit from particular medications [49], to determine whether to discontinue medications [50], and to assess the impacts of drug switching [51]. Future multimodal AI methods are anticipated to develop optimal dose prescriptions tailored to each patient, thus facilitating personalized medicine, similar to advancements already seen in the field of oncology [52].

Estimating recurrence is also critical for staging and treatment planning for discharged patients with CVD. A recent study employed a transformer-based algorithm, integrating surface ECG signals and clinical features, to significantly enhance predictive performance for atrial fibrillation (AF) recurrence postablation, superior to analyses of either modality alone [53]. Data from pre-ablation pulmonary vein computed tomography (PVCT) images also show good predictive performance for 1-year AF recurrence post-ablation [54]. Together, they reveal better predictive ability compared with traditional models (Fig. 4).

Proposing multimodality-AI models in building and optimizing operational platforms

Utilizing real-time data to enhance perioperative assessment

Within multimodal AI methods, various real-time data can be captured precisely, which is conducive to preoperative assessments. Preoperative assessments involving DL significantly contribute to a profound understanding of complex structures. Automated interpretation of images from invasive optical coherence tomography and intravascular ultrasound [55], and real-time data from transesophageal echocardiography enables full guidance for 3D reconstruction of detailed anatomy for coronary vessel interventions and structural repairments. In addition, noninvasive methods like CTA can provide 3D reconstruction of varying degrees of calcification, calculate fractional flow reserve, and monitor hemodynamics [56, 57]. Building on these advancements, the multimodality image-based model, involves the amalgamation of morphological and biomechanical factors at the same time. Models built from CTA and adjusted using invasive coronary angiography can provide predictions of acute coronary syndrome culprit lesions over conventional analysis before surgery [58]. Additionally. while using CTA and echocardiography, models based on hemodynamics of the valve leaflet motion and the resulting blood flow with detailed 3D patient geometry are applied for Transcatheter Aortic Valve Replacement (RAVR) device optimization [59]. These advancements reveal that the amalgamation of morphological and biomechanical factors simultaneously can provide more detailed structural information procedural processes.

The combination of computational simulations (CS), extended reality (XR), and multimodal AI together [60] could form a big AI operational platform to extend the previously mentioned holographic display of patients' heart anatomy at diagnosis into each stage of treatment. AI offers algorithms for decision-making, while XR is leveraged for immersive evaluations, extending the holographic display of patients' hearts into reality using CS. Furthermore, this synergistic integration of technologies will allow clinicians to provide ongoing feedback with predictive models to simulate various treatment scenarios and the potential outcomes of different operation plans. This will enhance decision-making accuracy and prompt personalized solutions. Moreover, with rising robotics [61], this big AI operational platform could revolutionize the whole process of operations from the perspectives of precise preoperative planning, recognition of operational phases, segmentation of anatomic markers, provision of detailed analytical reports for clinical physicians at each stage, as well as evaluations of cardiac function and the best method for patient prognosis with real-time data. Several platform prototypes are on the horizon, such as CARES Copilot 1.0 and SentiAR [62]. In the future, more and more similar platforms will thrive (Fig. 5).

Supporting cardiovascular materials and 3D printing with multimodal data

Beyond that, multimodal AI could help enhance the increasing use of 3D printing, bridging the gap in our understanding of the anatomic and physiological relationships in the human body [63]. 3D printing is based on precise mimicry of the deformation of both the implanted device and local tissue [64], requiring high-spatial resolution for remodeling. The current single-modal technique presents an incomprehensive picture of the native structures, hindering complete emulation of the heart itself [65]. Multimodal AI can meet this requirement by integrating the biomechanical properties of both functional and static states across different modalities. It allows for labeling of specific structures (such as the sinoatrial and atrioventricular nodes) and selecting optimal views of target structures, while simultaneously mimicking the deformation of both the implanted device and local tissue [64]. It can further help to address the limitations of 3D printing in replicating the dynamic physics of tissue deformation

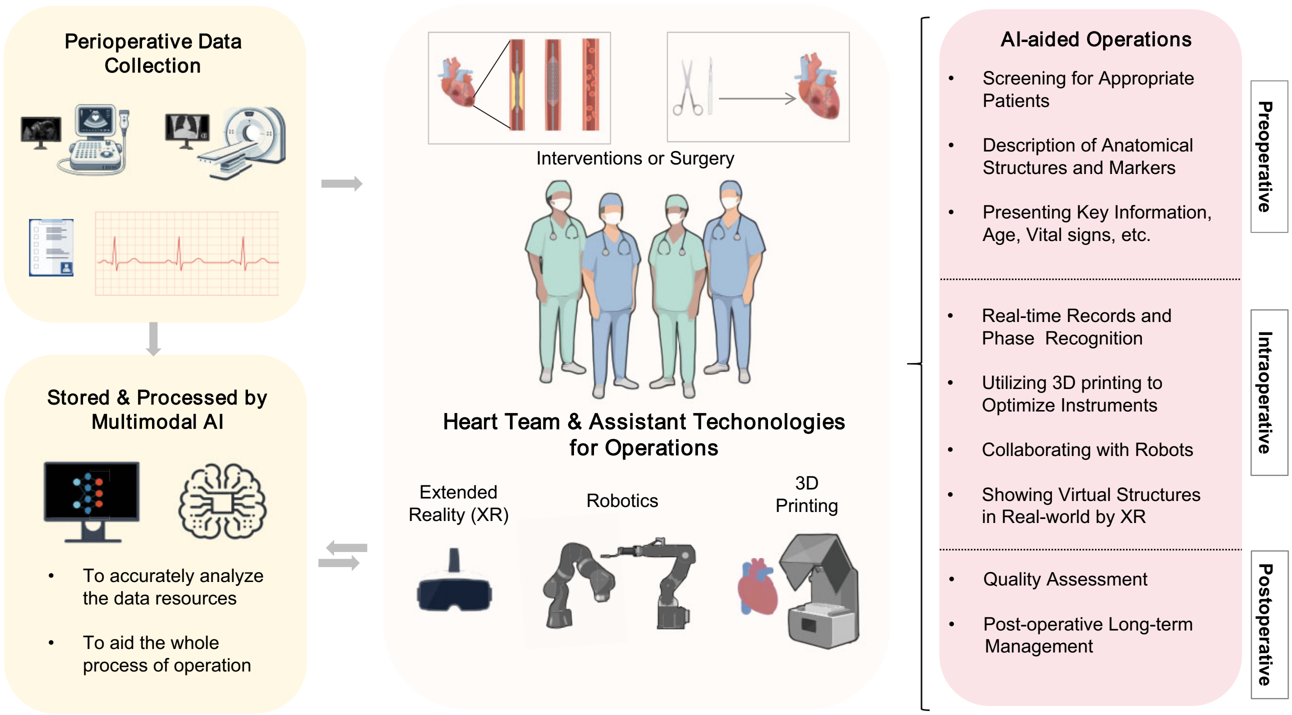


Figure 5. Multimodal AI platforms for operations. Perioperative data are first collected and then stored and processed by multimodal AI to analyze resources and aid the entire surgical process. With assistant technologies such as extended reality (XR), robotics and 3D printing, multimodal AI can aid the whole process of operations for the heart team, including surgery and interventions.

and flow, and the biomechanical properties of cardiac tissues [66] (Fig. 5).

Providing patient education and home-based/remote service with continuous monitoring

Despite the availability of best practices and treatments that are well understood in clinical settings, home-based cardiac rehabilitation and remote monitoring are emerging as a critical alternative to continuous monitoring and adjustment of treatments, potentially preventing disease progression and improving quality of life [67]. Cardiometabolic risk factor optimization is a long-term endeavor that ideally starts early in life and is sustained over time. Various standalone technologies, including smartphone apps, wearables, and in-home speakers are utilized to record daily life activities [68], offering cost-effective and continuous analysis, and identifying variables less prominently featured in traditional statistical approaches [69]. Combined with DL, they can synthesize different signals from people's daily life and habits to obtain information related to CVD risk factors [70, 71], alerting patients and providers to timely health status changes and the need for condition adjustments [72]. This could be done through online chatbots. By providing personalized coaching through voice assistance and text messages [73], they improve patient adherence to treatment plans by providing immediate responses, uninterrupted support, and understanding of conditions by answering questions and offering explanations [74]. Conjointly, these personalized approaches foster a collaborative environment where patients feel empowered to take charge of their health, thus improving clinical outcomes with better healthcare delivery and reducing hospitalizations to complete a significant shift toward patient-centered care out of hospital (Fig. 4).

Future perspectives

Throughout the field of cardiology, evidence is now mounting that demonstrates the advantages AI can offer, as discussed above. Remarkable performance has been witnessed in numerous clinically relevant tasks, tackling cumbersome challenges despite frequently being unimodal. With different values attached to each single modality, multimodal AI can flexibly evaluate and incorporate data into comprehensive responses. It is reported that within the next 5 years using today's technologies could result in savings of 5% to 10% of healthcare spending by improving clinical operations, quality and safety, and continuity of care [75]. Moreover, having acquired the most relevant information across modalities through the training process, medical knowledge could then be transferred from one modality to another for the generation of content or to make predictions, replacing harmful and invasive examinations by utilizing their relationships with non-invasive ones. More importantly, with interconnections being unearthed by multimodal AI, missed information may be complemented through another modality, contributing to solving the problems of medical fragmentation [76], and reducing medical errors and waste of medical resources caused by diffuse visits to physicians. In addition, when applied to clinical trials, it could help in propensity score matching [77] and finding comparable cohorts [78], adding up to building digital twin models in cardiovascular research [79].

However, advancements in multimodal biomedical AI are not devoid of challenges. We still lack the strength and specificity of models for widespread and independent adoption. Of primary concern is data quality and quantity. Considering the vastness and diversity of multimodal healthcare data, there is a need for enhanced techniques to collect and standardize the data. Plus, current healthcare data exhibit biases, showing a lack of ethnic diversity and an imbalance in socioeconomic representation, fur-

Table 4. Algorithms used for analyzing different modality combinations in AI for cardiovascular disease.

Modality involved	Analytical algorithm	Advantages	Disadvantages	Representative article	Year	Summary of article	Performance
Imaging + imaging							
CT + MRI	Multiple learning network	End-to-end learning between images and meshes for modeling applications; directly support different cardiac simulations without additional efforts.	Training instability; difficulty in controlling output	Yu *et al.* [18]	2021	Combing multiple neural networks, based on adversarial reverse mapping for estimating multitype cardiac indices in different imaging modalities, e.g. MRI and CT, explores task dependencies by learning the mapping from multitype cardiac indices to cardiac images via adversarial training.	Surpassing existing single-modality multitask learning methods and achieving cross-modal knowledge transfer.
CT + MRI	GNN + CNN	Simulation-ready whole heart meshes from cardiac image data allows switching of template meshes to accommodate different modeling requirements	Computationally intensive; requires significant tuning and expertise	Kong *et al.* [21]	2023	Constructing meshes from 3D patient images by learning to deform a small set of deformation handles on a whole heart template.	Outperforming prior state-of-the-art methods and able to produce better geometries for cardiac flow simulations.
Echocardiography + MRI	CNN	Independently predict the development and progression of AS, regardless of traditional Doppler parameters and baseline AS severity.	Cardiovascular phenotype that is not exclusive to AS, a condition that shares several risk factors with CAD, stroke, and HF	Oikonomou *et al.* [43]	2024	A video-based AI biomarker can detect severe AS using single-view long-axis echocardiography without Doppler characterization.	AUC: 0.70 and 0.67 in two separated cohorts
Echocardiography + CT	CNN	The meshes generated by this model significantly outperform traditional methods in terms of spatial accuracy and element quality, particularly excelling in the reconstruction of the left ventricle, aorta, and aortic valve.	Still limited on complex structures, such as in congenital diseases	Ozturk *et al.* [59]	2025	A start-to-end computational pipeline for patient-specific hemodynamic analysis for AS using CT imaging patient data.	NA

Table 4. (*Continued*)

Modality involved	Analytical algorithm	Advantages	Disadvantages	Representative article	Year	Summary of article	Performance
CTA + ICA	CNN	Significantly enhanced the predictive ability for culprit lesions of acute coronary syndrome.	Retrospective data and partially omitting information.	Koo *et al.* [58]	2024	The current study investigated the best predictors among AI-enabled coronary CTA-derived quantitative plaque and hemodynamic features in the prediction of acute coronary syndrome culprit lesions and their incremental prognostic value to the conventional coronary CTA analysis.	AUC: 0.86 for derivation cohort, 0.84 for validation cohort
Imaging + omics							
CTA + Omics	CNN + ML classifiers	Multi-omics, radiomics, and immunological markers complement traditional clinical risk factors.	Small sample size and single-center	Zhang *et al.* [34]	2024	Two distinct AAA subtypes were identified for risk stratification and CXCL1 overexpression activated neutrophils through the NF-κB pathway, contributing to AAA development.	AUC for subtype prediction: 0.927
MRI + omics	FCN	New modality combination enables identification of genetic variants for low absolute event rates, which may enhance risk stratification beyond familial disease.	Biased population	De Marvao *et al.* [36]	2021	Combining whole exome sequencing and cardiac magnetic resonance to identify specific genes that are relative with higher risk in patients with HCM.	In this population of middle-aged adults, SARC-HCM-P/LP variants have low aggregate penetrance for overt HCM but are associated with an increased risk of adverse cardiovascular outcomes and an attenuated cardiomyopathic phenotype.
MRI + omics	CNN	Capturing cardiac dynamic flow volumes from phase-contrast cardiac MRI data at scale.	Without manually curating every imaging study some error may be introduced	Gomes *et al.* [24]	2024	Combining deep learning strategies with available cardiac MRI and GWAS with data from the UK Biobank to reveal plausible causal relationships among key cardiac dynamic flow volume parameters using genetic instruments.	NA

Table 4. (*Continued*)

Modality involved	Analytical algorithm	Advantages	Disadvantages	Representative article	Year	Summary of article	Performance
Multiomics							
Multiomics	8 ML classifiers	By obtaining a complete set of omics for patients, it provides a straightforward example that this multi-omics method shows strong classification performance, excelling in unambiguous diagnosis of the major subtypes of EHT.	Small samples	Reel *et al.* [35]	2022	They train ML algorithms for diagnosing endocrine hypertension subtypes using multiomics data and provide an understanding of discriminating features and their importance to different disease combinations.	AUC: 0.96; sensitivity: 90%; specificity: 86%
ECG + imaging							
ECG + SPECT	Different neural network + ML classifiers	The study combines ECG, gated myocardial perfusion SPECT(GMPS), and clinical variables to predict CRT response, which is a novel approach.	Small population	De A Fernandes *et al.* [48]	2023	Compared to guideline criteria, ML methods trended toward improved CRT response and super-response prediction.	AUC: 0.80; sensitivity: 0.86; specificity: 0.75
ECG + echocardio-graphy	Different CNN algorithms	Combining complementary knowledge from multiple modalities to improve diagnostic performance. Different spatiotemporal convolutions for video-based classification winning over other models when combining echocardiography with ECG.	Slower training times and complicated hypermeters adjusting	Soto *et al.* [28]	2022	Using multimodal nueral networks to analyze combined instances of electrocardiograms and echocardiograms from 2728 patients to distinguish individuals with HCM from individuals with hypertension (HTN).	F1 score: 0.71 for HCM, 0.96 for HTN
ECG + X-ray	CNN	The model was significantly more accurate than the conventional tree algorithm.	Small population and simplified classification	Nishimori *et al.* [23]	2021	They assessed the efficacy of the AI model using ECG and chest X-rays to identify the location of cardiac accessory pathways.	F1-score: 0.88; sensitivity: 0.99

Table 4. (*Continued*)

Modality involved	Analytical algorithm	Advantages	Disadvantages	Representative article	Year	Summary of article	Performance
EHR + imaging/ECG							
EHR + CT	Different CNN algorithms + ML classifiers	The integration of I-Score is capable of selecting the most influential combination of variables from a large number of variables, further analyzing the correlation between different variable combinations and labels, and enhancing the predictive performance of the model.	Disparities in expertise of electrophysiologists may influence the results	Liu *et al.* [54]	2024	I-Score combined with featured clinical variables can predict recurrence in paroxysmal AF patients before catheter ablation. Application of this prediction model in AF recurrence can classify patients with high risk for AF recurrence who are required for post-ablation close follow-up.	AUC: 0.76; sensitivity: 86.7%; specificity: 51.0%
EHR + echocardiography	Multiple ML classifiers	The risk score outperformed other scores in predicting 1-year mortality after M-TEER.	Population biased	Hausleiter *et al.* [46]	2024	To develop and validate a comprehensive risk score for the prediction of mortality and clinical outcomes from the large extended patient population.	AUC: 0.789
EHR + ECG	Transformer	Focusing on the temporal continuity of the data and captures the hidden deep features well.	Lack of clear interpretability	Qiu *et al.* [53]	2024	A novel algorithm based on Transformer using surface ECG signals and clinical features can predict AF recurrence.	Sensitivity: 81.1%; specificity: 81.7%; AUC: 0.899; F1 score: 71.7%
Other combinations							
EHR + medical examinations	CNN + RNN	A two-step training algorithm was proposed, which first involves pre-training and then fine-tuning, enhancing the training efficiency of the model.	Complexity and overfitting; the limitation of dataset	Mousavi *et al.* [72]	2020	Applying CNN to automatically extract time-invariant features, and attention mechanism to put more emphasis on the important regions of the segmented input signal(s) together to suppress the false alarms in ICUs, especially arrhythmia.	Sensitivity: 93.88%; specificity: 92.05%; AUC: 0.93

Table 4. (*Continued*)

Modality involved	Analytical algorithm	Advantages	Disadvantages	Representative article	Year	Summary of article	Performance
Face videos + EHR	CNN	Novel data combination.	Tested on well-controlled small-scale databases instead of real-life situations	Niu *et al.* [71]	2018	Remote HR estimation, which is a large-scale multi-modal database recorded with various head movement, illumination variations, and acquisition device changes.	Pearsons correlation coefficient r: 0.73
Remote devices + EHR	FNN + ML classifiers	Complementing and expanding upon variables of interest by identifying time-dependent changes.	The positive predictive value of these models may underestimate model performance in a population at greater risk due to low patient enrollment	Ginder *et al.* [69]	2023	Evaluating whether daily remote-monitoring data may predict appropriate ICD therapies for ventricular tachycardia or ventricular fibrillation.	Neural network Sensitivity: 54%; specificity: 96%; AUC: 0.90 ML classifiers Sensitivity: 39%; specificity: 91%; AUC: 0.72
Admission notes and clinical tabular data	BERT	Multimodality could further enhance the model's ability and credibility to evaluate outcomes compared to unimodality.	Omitting recorded data during the treatment and simplifying the feature extraction	Gao *et al.* [44]	2024	The multimodal DL model for combining admission notes and clinical tabular data showed promising efficacy as a potentially novel method in evaluating the risk of mortality in patients with HF, providing more accurate and timely decision support.	AUC: 0.838 for the internal validation sets, 0.767 for the external validation sets.
Combination >2 modalities							
EHR + activities + wearable devices.	Neural network + different ML classifiers	Multimodal physiological data, collected in a non-laboratory environment, are able to predict the relative intensity category. The examination of each of these modalities can be independent or combined (using fusion).	Biased data	Chowdhury *et al.* [70]	2019		AUC highest: 85.2%

Table 4. (*Continued*)

Modality involved	Analytical algorithm	Advantages	Disadvantages	Representative article	Year	Summary of article	Performance
CT + SPECT + EHR	RNN + ML classifiers	It could offer additional information about HF risk in individuals who have already undergone SPECT without the need for further testing and identifying patients at high risk for HF exacerbation.	Using high-spatial-resolution MRI, which may limit the model generalization	Feher *et al.* [45]	2024	It aimed to assess whether AI models incorporating clinical, stress test, and imaging parameters could predict hospitalization for acute HF exacerbation in patients undergoing SPECT/CT myocardial perfusion imaging.	AUC: 0.87 ± 0.03 for internal validation, 0.80 ± 0.04 for external validation
X-Ray + EHR + clinical parameters	Transformer	Improving performance with multimodal data, handling missing data, coping wirh extensive data, and allowing insight into the network's decision-making process.	Developed in supervised learning, requiring labels for each patient, limiting the range of data used Domain transferability of model not tested	Khader *et al.* [30]	2023	To develop a DL algorithm capable of integrating multimodal patient data and compare its performance to models incorporating a single modality for diagnosing up to 25 pathologic conditions.	AUC: 0.77

ML classifiers in the table refer to algorithms as: *k*-means, support vector machine (SVM), random forest (RF), etc. GNN, graph neural network; FNN, feedforward neural network; RNN, recurrent neural network; BERT, bidirectional encoder representations from transformers; HTN, occult hypertension; EHT, endocrine hypertension; GMPS, gated myocardial perfusion SPECT; CRT, cardiac resynchronization therapy; M-TEER, mitral valve TEER; GWAS, genome-wide association study; ICD, implantable cardioverter defibrillator; NA, not applicable.

ther complicated by disparities in resource and technological access [80]. Imbalances in training data cause AI algorithms to perform heterogeneously across different subpopulations. A pneumothorax algorithm accurate at one hospital may lose >10% accuracy at another [81]. Another example is the pulse oximeter. The accuracy of oxygen saturation varies with skin tone because infrared signals interact with skin pigmentation, which may overestimate Saturation of peripheral oxygen (SpO_2) at low Saturation of arterial oxygen (SaO_2) in dark-skinned patients [82], risking inadequate oxygen therapy and vital-organ damage. Algorithmic tools can only be as good as the devices feeding data into them. ML from biased data yields biased findings [83], calling for more efforts to contribute to generating standardized and consistent data of high quality, moving beyond semantic differences, temporal variability, and ethical constraints. Big data platforms, including large-scale cohort studies [84], consortia, and omics databases [85] are among the best approaches to obtaining high-standard and cross-scale data.

Moreover, as examination methods and modality types increase, experienced experts with training in anomaly detection and advanced modal fusion techniques for lifelong learning are necessary [86]. Common combinations have been witnessed (Table 4), yet new modalities are springing up. For instance, how to integrate self-reporting sensors accompanying implantable devices (stents, valves, etc.) *in vivo* and from lightweight wearable devices poses a challenge for dynamic monitoring [87]. This places higher requirements on clinical doctors' comprehension and processing skills regarding AI.

Simultaneously, an increase in the dimensionality of data brings out another important problem: model interpretability and explainability. Comprising millions of parameters and hundreds of layers, the very complexity of AI tools renders them as opaque 'black-boxes' [88], as the computational mechanisms between input and output are too intricate to afford human understanding. Without understanding how AI arrives at outputs, it is hard for healthcare professionals and patients to trust AI tools [89, 90]: who is accountable if AI-driven medical diagnostics or treatments fail? Can we obtain an explainable AI diagnosis of why the failure occurred [91, 92]? To address this, one way is to start by choosing simpler, interpretable models like decision trees when suitable. Yet to ensure model complexity, multiple techniques have been proposed to solve this problem, including building locally linear models around the predictions of an opaque model for explanations (LIME) [93], allocating feature contribution to predictions using game theory (SHAP) [94], and highlighting the important regions in the image for predicting the concept in CNN (Grad-CAM) [95]. Nonetheless, most of the explaining methods are unimodal. The introduction of multimodal data significantly intensifies the difficulty of illustration, and aligning semantics across different data types may make model comprehension even more challenging. Additionally, no unified protocol exists to verify if explanations accurately reflect cross-modal correlations. For instance, in developing multimodal AI systems for disease diagnosis (integrating imaging, genetic data, and clinical records), standardized metrics are lacking to evaluate explanation quality. The complex decision-making process obscures traceability to specific inputs [96] and medical experts' subjective assessments of explanations may vary based on individual expertise. Therefore, more complex and precise explanation models will become imperative.

To promote the development of this field, experts globally should be organized to discuss the establishment and standards of building multimodal AI tools, addressing problems of inconsistent conceptual understanding among professionals, as well as unreasonable or novel data inclusion and classification. General consensus and reporting guidelines will ensure the secure storage, processing, and sharing of data, paving the way for promoting future interdisciplinary collaboration among different organizations [97]. Meanwhile, mandatory legislation and policies from the government also serve as prerequisites to guarantee the transparency and data privacy of AI technology [98]. Therefore, federated and nationwide organizations of clinicians, technology developers, and the public, in cooperation with the government, can help build a prescription of intended workflows with standardized processes, achieving the speedy adoption of useful AI technologies to protect patient well-being.

Conclusion

In the coming decade, AI technologies will become an essential part of the cardiovascular diagnostic, prognostic, and therapeutic toolkit, with multimodality becoming more and more predominant. They are poised to enable more detailed phenotyping, reshape disease classifications, facilitate biomarker discovery, reduce the risks of developing new therapies, as well as support more wearable, portable, real-time, and continuous monitoring. As technological advancements continue to outpace regulatory adaptation, it is also crucial to integrate ethical, legitimate, and trustworthy AI into the development and validation stages. More efforts should be made to ensure the effective application of AI in the real-world of dynamic environments.

Acknowledgments

Figures 4 and 5 were created with Figdraw image and illustration software.

Author contributions

X.Y., Y.L., and J.W. wrote the original article, X.Y. and M.C. conceived the original ideas, M.C and Z.Y. supervised the manuscript, X.Y. and Y.J. completed visualization, M.C. and Y.L. helped with revision and editing, M.C. and J.W. completed funding acquisition. All authors read and approved the final manuscript and participated in final approval of the version to be published.

Conflict of interest

None declared.

Funding

This work was supported by the National Natural Science Foundation of China (grant Nos. 62306192, U23A20395, 82170375, and 62476185), 1.3.5 project for disciplines of excellence from West China Hospital of Sichuan University (grant Nos. ZYGD23021 and 23HXFH009), and the Natural Science Foundation of Sichuan Province, China (grant No. 2023NSFSC1638).

References

1. Lu MY, Williamson DFK, Chen TY *et al.* Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat Biomed Eng* 2021;**5**:555–70. https://doi.org/10.1038/s41551-020-00682-w.
2. Muse ED, Topol EJ. Transforming the cardiometabolic disease landscape: multimodal AI-powered approaches in prevention

and management. *Cell Metab* 2024;**36**:670–83. https://doi.org/10.1016/j.cmet.2024.02.002.

3. Jaltotage B, Lu J, Dwivedi G. Use of artificial intelligence including multimodal systems to improve the management of cardiovascular disease. *Can J Cardiol* 2024;**40**:1804–12. https://doi.org/10.1016/j.cjca.2024.07.014.
4. van der Hoeven BL, Schalij MJ, Delgado V. Multimodality imaging in interventional cardiology. *Nat Rev Cardiol* 2012;**9**:333–46. https://doi.org/10.1038/nrcardio.2012.14.
5. Berman DS. Fourth annual Mario S. Verani, MD Memorial Lecture: noninvasive imaging in coronary artery disease: changing roles, changing players. *J Nucl Cardiol* 2006;**13**:457–73. https://doi.org/10.1016/j.nuclcard.2006.05.009.
6. Jone P-N, Haak A, Petri N *et al.* Echocardiography-fluoroscopy fusion imaging for guidance of congenital and structural heart disease interventions. *JACC: Cardiovasc Imaging* 2019;**12**:1279–82. https://doi.org/10.1016/j.jcmg.2018.11.010.
7. Gao J, Li P, Chen Z *et al.* A survey on deep learning for multimodal Data fusion. *Neural Comput* 2020;**32**:829–64. https://doi.org/10.1162/neco_a_01273.
8. Acosta JN, Falcone GJ, Rajpurkar P *et al.* Multimodal biomedical AI. *Nat Med* 2022;**28**:1773–84. https://doi.org/10.1038/s41591-022-01981-2.
9. Azam MA, Khan KB, Salahuddin S *et al.* A review on multimodal medical image fusion: compendious analysis of medical modalities, multimodal databases, fusion techniques and quality metrics. *Comput Biol Med* 2022;**144**:105253. https://doi.org/10.1016/j.compbiomed.2022.105253.
10. Ngiam J, Khosla A, Kim M *et al.* Multimodal deep learning. *In Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, 2011, p.689–96.
11. Li J, Li D, Savarese S *et al.* BLIP-2: bootstrapping language-image pre-training with frozen image encoders and large language models. In *Proceedings of the 40th International Conference on Machine Learning*. PMLR, 2023, p.19730–42.
12. Lu MY, Chen B, Williamson DF *et al.* A multimodal generative AI copilot for human pathology. *Nature* 2024;**634**:466–73. https://doi.org/10.1038/s41586-024-07618-3.
13. Christensen M, Vukadinovic M, Yuan N *et al.* Vision–language foundation model for echocardiogram interpretation. *Nat Med* 2024;**30**:1481–8. https://doi.org/10.1038/s41591-024-02959-y.
14. Zhou J, He X, Sun L *et al.* Pre-trained multimodal large language model enhances dermatological diagnosis using SkinGPT-4. *Nat Commun* 2024;**15**:5649. https://doi.org/10.1038/s41467-024-50043-3.
15. Zhao T, Gu Y, Yang J *et al.* A foundation model for joint segmentation, detection and recognition of biomedical objects across nine modalities. *Nat Methods* 2025;**22**:166–76. https://doi.org/10.1038/s41592-024-02499-w.
16. Baltrusaitis T, Ahuja C, Morency L-P. Multimodal machine learning: A survey and taxonomy. *IEEE Trans Pattern Anal Mach Intell* 2019;**41**:423–43. https://doi.org/10.1109/TPAMI.2018.2798607.
17. Zhao F, Zhang C, Geng B. Deep multimodal data fusion. *ACM Comput Surv* 2024;**56**:1–36. https://doi.org/10.1145/3649447.
18. Yu C, Gao Z, Zhang W *et al.* Multitask learning for estimating multitype cardiac indices in MRI and CT based on adversarial reverse mapping. *IEEE Transactions on Neural Networks and Learning Systems* 2021;**32**:493–506. https://doi.org/10.1109/TNNLS.2020.2984955.
19. Kong F, Wilson N, Shadden S. A deep-learning approach for direct whole-heart mesh reconstruction. *Med Image Anal* 2021;**74**:102222. https://doi.org/10.1016/j.media.2021.102222.
20. Ta K, Ahn SS, Thorn SL *et al.* Multi-task learning for motion analysis and segmentation in 3D echocardiography. *IEEE Trans Med Imaging* 2024;**43**:2010–20. https://doi.org/10.1109/TMI.2024.3355383.
21. Kong F, Shadden SC. Learning whole heart mesh generation from patient images for computational simulations. *IEEE Trans Med Imag* 2023;**42**:533–45. https://doi.org/10.1109/TMI.2022.3219284.
22. Laumer F, Amrani M, Manduchi L *et al.* Weakly supervised inference of personalized heart meshes based on echocardiography videos. *Med Image Anal* 2023;**83**:102653. https://doi.org/10.1016/j.media.2022.102653.
23. Nishimori M, Kiuchi K, Nishimura K *et al.* Accessory pathway analysis using a multimodal deep learning model. *Sci Rep* 2021;**11**:8045. https://doi.org/10.1038/s41598-021-87631-y.
24. Gomes B, Singh A, O'Sullivan JW *et al.* Genetic architecture of cardiac dynamic flow volumes. *Nat Genet* 2024;**56**:245–57. https://doi.org/10.1038/s41588-023-01587-5.
25. Pirruccello JP, Chaffin MD, Chou EL *et al.* Deep learning enables genetic analysis of the human thoracic aorta. *Nat Genet* 2022;**54**:40–51. https://doi.org/10.1038/s41588-021-00962-4.
26. Tabassum R, Rämö JT, Ripatti P *et al.* Genetic architecture of human plasma lipidome and its link to cardiovascular disease. *Nat Commun* 2019;**10**:4329. https://doi.org/10.1038/s41467-019-11954-8.
27. Shah M, De A, Inácio MH *et al.* Environmental and genetic predictors of human cardiovascular ageing. *Nat Commun* 2023;**14**:4941. https://doi.org/10.1038/s41467-023-40566-6.
28. Soto JT, Weston Hughes J, Sanchez PA *et al.* Multimodal deep learning enhances diagnostic precision in left ventricular hypertrophy. *European Heart Journal—Digital Health* 2022;**3**:380–9. https://doi.org/10.1093/ehjdh/ztac033.
29. Tang P, Yan X, Nan Y *et al.* FusionM4Net: A multi-stage multimodal learning algorithm for multi-label skin lesion classification. *Med Image Anal* 2022;**76**:102307. https://doi.org/10.1016/j.media.2021.102307.
30. Khader F, Müller-Franzes G, Wang T *et al.* Multimodal deep learning for integrating chest radiographs and clinical parameters: A case for transformers. *Radiology* 2023;**309**:e230806. https://doi.org/10.1148/radiol.230806.
31. Maiorino E, Loscalzo J. Phenomics and robust multiomics data for cardiovascular disease subtyping. *Arterioscler Thromb Vasc Biol* 2023;**43**:1111–23. https://doi.org/10.1161/ATVBAHA.122.318892.
32. Banerjee A, Dashtban A, Chen S *et al.* Identifying subtypes of heart failure from three electronic health record sources with machine learning: an external, prognostic, and genetic validation study. *The Lancet Digital Health* 2023;**5**:e370–9. https://doi.org/10.1016/S2589-7500(23)00065-1.
33. Karczewski KJ, Snyder MP. Integrative omics for health and disease. *Nat Rev Genet* 2018;**19**:299–310. https://doi.org/10.1038/nrg.2018.4.
34. Zhang L, Yang H, Zhou C *et al.* Artificial intelligence-driven multiomics predictive model for abdominal aortic aneurysm subtypes to identify heterogeneous immune cell infiltration and predict disease progression. *Int Immunopharmacol* 2024;**138**:112608. https://doi.org/10.1016/j.intimp.2024.112608.
35. Reel PS, Reel S, van Kralingen JC *et al.* Machine learning for classification of hypertension subtypes using multi-omics:

A multi-centre, retrospective, data-driven study. *EBioMedicine* 2022;**84**:104276. https://doi.org/10.1016/j.ebiom.2022.104276.
36. De Marvao A, McGurk KA, Zheng SL *et al.* Phenotypic expression and outcomes in individuals with rare genetic variants of hypertrophic cardiomyopathy. *J Am Coll Cardiol* 2021;**78**:1097–110. https://doi.org/10.1016/j.jacc.2021.07.017.
37. Lloyd-Jones DM, Allen NB, Anderson CAM *et al.* Life's Essential 8: updating and enhancing the American Heart Association's Construct of Cardiovascular Health: A presidential advisory from the American Heart Association. *Circulation* 2022;**146**:e18–43.. https://doi.org/10.1161/CIR.0000000000001078.
38. Grundy SM, Stone NJ, Bailey AL *et al.* 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the Management of blood Cholesterol: A report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation* 2019;**139**:e1082–143. https://doi.org/10.1161/CIR.0000000000000625.
39. Yang X, Li J, Hu D *et al.* Predicting the 10-year risks of atherosclerotic cardiovascular disease in Chinese population: the China-PAR Project (Prediction for ASCVD Risk in China). *Circulation* 2016;**134**:1430–40. https://doi.org/10.1161/CIRCULATIONAHA.116.022367.
40. Polonsky TS, Greenland P. CVD screening in low-risk, asymptomatic adults: clinical trials needed. *Nat Rev Cardiol* 2012;**9**:599–604. https://doi.org/10.1038/nrcardio.2012.114.
41. Clapp MA, Kim E, James KE *et al.* Comparison of natural language processing of clinical notes with a validated risk-stratification tool to predict severe maternal morbidity. *JAMA Netw Open* 2022;**5**:e2234924. https://doi.org/10.1001/jamanetworkopen.2022.34924.
42. Forrest IS, Petrazzini BO, Duffy Á *et al.* Machine learning-based marker for coronary artery disease: derivation and validation in two longitudinal cohorts. *Lancet (London, England)* 2023;**401**:215–25. https://doi.org/10.1016/S0140-6736(22)02079-7.
43. Oikonomou EK, Holste G, Yuan N *et al.* A multimodal video-based AI biomarker for aortic stenosis development and progression. *JAMA Cardiol* 2024;**9**:534–44. https://doi.org/10.1001/jamacardio.2024.0595.
44. Gao Z, Liu X, Kang Y *et al.* Improving the prognostic evaluation precision of hospital outcomes for heart failure using admission notes and clinical tabular data: multimodal deep learning model. *J Med Internet Res* 2024;**26**:e54363. https://doi.org/10.2196/54363.
45. Feher A, Bednarski B, Miller RJ *et al.* Artificial intelligence predicts hospitalization for acute heart failure exacerbation in patients undergoing myocardial perfusion imaging. *J Nucl Med* 2024;**65**:768–74. https://doi.org/10.2967/jnumed.123.266761.
46. Hausleiter J, Lachmann M, Stolz L *et al.* Artificial intelligence-derived risk score for mortality in secondary mitral regurgitation treated by transcatheter edge-to-edge repair: the EuroSMR risk score. *Eur Heart J* 2024;**45**:922–36. https://doi.org/10.1093/eurheartj/ehad871.
47. Gautam N, Ghanta SN, Clausen A *et al.* Contemporary applications of machine learning for device therapy in heart failure. *JACC Heart Failure* 2022;**10**:603–22. https://doi.org/10.1016/j.jchf.2022.06.011.
48. de A Fernandes F, Larsen K, He Z *et al.* A machine learning method integrating ECG and gated SPECT for cardiac resynchronization therapy decision support. *Eur J Nucl Med Mol Imaging* 2023;**50**:3022–33. https://doi.org/10.1007/s00259-023-06259-4.
49. Chowdhury S, Chen Y, Li P *et al.* Stratifying heart failure patients with graph neural network and transformer using Electronic Health Records to optimize drug response prediction. *J Am Med Inform Assoc* 2024;**31**:1671–81. https://doi.org/10.1093/jamia/ocae137.
50. Mora D, Nieto JA, Mateo J *et al.* Machine learning to predict outcomes in patients with acute pulmonary embolism who prematurely discontinued anticoagulant therapy. *Thromb Haemost* 2022;**122**:570–7. https://doi.org/10.1055/a-1525-7220.
51. Mele M, Mele A, Imbrici P *et al.* Pleiotropic effects of direct oral anticoagulants in chronic heart failure and atrial fibrillation: machine learning analysis. *Molecules* 2024;**29**:2651. https://doi.org/10.3390/molecules29112651.
52. Wu X, Li W, Tu H. Big data and artificial intelligence in cancer research. *Trends Cancer* 2024;**10**:147–60. https://doi.org/10.1016/j.trecan.2023.10.006.
53. Qiu Y, Guo H, Wang S *et al.* Deep learning-based multimodal fusion of the surface ECG and clinical features in prediction of atrial fibrillation recurrence following catheter ablation. *BMC Med Inf Decis Making* 2024;**24**:225. https://doi.org/10.1186/s12911-024-02616-x.
54. Liu C-M, Chen W-S, Chang S-L *et al.* Use of artificial intelligence and I-score for prediction of recurrence before catheter ablation of atrial fibrillation. *Int J Cardiol* 2024;**402**:131851. https://doi.org/10.1016/j.ijcard.2024.131851.
55. Yang S, Kweon J, Roh J-H *et al.* Deep learning segmentation of major vessels in X-ray coronary angiography. *Sci Rep* 2019;**9**:16897. https://doi.org/10.1038/s41598-019-53254-7.
56. Griffin WF, Choi AD, Riess JS *et al.* AI evaluation of stenosis on coronary CTA, comparison with quantitative coronary angiography and fractional flow reserve. *JACC: Cardiovasc Imaging* 2023;**16**:193–205. https://doi.org/10.1016/j.jcmg.2021.10.020.
57. Yi Y, Xu C, Xu M *et al.* Diagnostic improvements of deep learning–Based image reconstruction for assessing calcification-related obstructive coronary artery disease. *Front Cardiovasc Med* 2021;**8**:758793. https://doi.org/10.3389/fcvm.2021.758793.
58. Koo B-K, Yang S, Jung JW *et al.* Artificial intelligence–Enabled quantitative coronary plaque and hemodynamic analysis for predicting acute coronary syndrome. *JACC: Cardiovasc Imaging* 2024;**17**:1062–76. https://doi.org/10.1016/j.jcmg.2024.03.015.
59. Ozturk C, Pak DH, Rosalia L *et al.* AI-powered multimodal modeling of personalized hemodynamics in aortic stenosis. *Adv Sci (Weinh)* 2025;**12**:e2404755. https://doi.org/10.1002/advs.202404755.
60. Chessa M, Van De Bruaene A, Farooqi K *et al.* Three-dimensional printing, holograms, computational modelling, and artificial intelligence for adult congenital heart disease care: an exciting future. *Eur Heart J* 2022;**43**:2672–84. https://doi.org/10.1093/eurheartj/ehac266.
61. Hashimoto DA, Rosman G, Rus D *et al.* Artificial intelligence in surgery: promises and perils. *Ann Surg* 2018;**268**:70–6. https://doi.org/10.1097/SLA.0000000000002693.
62. Sentiar AR.https://sentiar.com(Accessed date: 10 May, 2025)
63. Wang DD, Qian Z, Vukicevic M *et al.* 3D Printing, computational modeling, and artificial intelligence for structural heart disease. *JACC Cardiovascular Imaging* 2021;**14**:41–60. https://doi.org/10.1016/j.jcmg.2019.12.022.
64. Vukicevic M, Mehta SM, Grande-Allen KJ *et al.* Development of 3D printed mitral valve constructs for transcatheter device modeling of tissue and device deformation. *Ann Biomed Eng* 2022;**50**:426–39. https://doi.org/10.1007/s10439-022-02927-y.

65. Vernon MJ, Mela P, Dilley RJ *et al.* 3D printing of heart valves. *Trends Biotechnol* 2024;**42**:612–30. https://doi.org/10.1016/j.tibtech.2023.11.001.
66. Vukicevic M, Mosadegh B, Min JK *et al.* Cardiac 3D printing and its future directions. *JACC Cardiovascular Imaging* 2017;**10**:171–84. https://doi.org/10.1016/j.jcmg.2016.12.001.
67. Wong CK, Hai JJ, Lau Y-M *et al.* Protocol for home-based solution for remote atrial fibrillation screening to prevent recurrence stroke (HUA-TUO AF Trial): a randomised controlled trial. *BMJ Open* 2022;**12**:e053466. https://doi.org/10.1136/bmjopen-2021-053466.
68. Cowie MR, Lam CSP. Remote monitoring and digital health tools in CVD management. *Nat Rev Cardiol* 2021;**18**:457–8. https://doi.org/10.1038/s41569-021-00548-x.
69. Ginder C, Li J, Halperin JL *et al.* Predicting malignant ventricular arrhythmias using real-time remote monitoring. *J Am Coll Cardiol* 2023;**81**:949–61. https://doi.org/10.1016/j.jacc.2022.12.024.
70. Chowdhury AK, Tjondronegoro D, Chandran V *et al.* Prediction of relative physical activity intensity using multimodal sensing of physiological data. *Sensors* 2019;**19**:4509. https://doi.org/10.3390/s19204509.
71. Niu X, Han H, Shan S *et al.* VIPL-HR: A multi-modal database for pulse estimation from less-constrained face video. 2018. https://doi.org/10.48550/arXiv.1810.04927.
72. Mousavi S, Fotoohinasab A, Afghah F. Single-modal and multi-modal false arrhythmia alarm reduction using attention-based convolutional and recurrent neural networks. *PLoS One* 2020;**15**:e0226990. https://doi.org/10.1371/journal.pone.0226990.
73. Hassoon A, Baig Y, Naiman DQ *et al.* Randomized trial of two artificial intelligence coaching interventions to increase physical activity in cancer survivors. *Npj Digit Med* 2021;**4**:168. https://doi.org/10.1038/s41746-021-00539-9.
74. Kozaily E, Geagea M, Akdogan ER *et al.* Accuracy and consistency of online large language model-based artificial intelligence chat platforms in answering patients' questions about heart failure. *Int J Cardiol* 2024;**408**:132115. https://doi.org/10.1016/j.ijcard.2024.132115.
75. Allen B, Agarwal S, Coombs L *et al.* 2020 ACR Data Science Institute Artificial Intelligence Survey. *Journal of the American College of Radiology: JACR* 2021;**18**:1153–9. https://doi.org/10.1016/j.jacr.2021.04.002.
76. Subbiah V. Fragmentation in medicine harms patients and hinders research. *Nat Med* 2024;**30**:2394–. https://doi.org/10.1038/s41591-024-03194-1.
77. Ghosh S, Boucher C, Bian J *et al.* Propensity score synthetic augmentation matching using generative adversarial networks (PSSAM-GAN). *Computer Methods and Programs in Biomedicine Update* 2021;**1**:100020. https://doi.org/10.1016/j.cmpbup.2021.100020.
78. Averitt AJ, Vanitchanant N, Ranganath R *et al.* The Counterfactual χ -GAN: finding comparable cohorts in observational health data. *J Biomed Inform* 2020;**109**:103515. https://doi.org/10.1016/j.jbi.2020.103515.
79. Corral-Acero J, Margara F, Marciniak M *et al.* The 'Digital Twin' to enable the vision of precision cardiology. *Eur Heart J* 2020;**41**:4556–64. https://doi.org/10.1093/eurheartj/ehaa159.
80. DeCamp M, Lindvall C. Mitigating bias in AI at the point of care. *Science* 2023;**381**:150–2. https://doi.org/10.1126/science.adh2713.
81. Leshem A, Segal E, Elinav E. The gut microbiome and individual-specific responses to diet. *mSystems* 2020;**5**:e00665–20. https://doi.org/10.1128/mSystems.00665-20.
82. Feiner JR, Severinghaus JW, Bickler PE. Dark skin decreases the accuracy of pulse oximeters at low oxygen saturation: the effects of oximeter probe type and gender. *Anesth Analg* 2007;**105**:S18–23. https://doi.org/10.1213/01.ane.0000285988.35174.d9.
83. Gianfrancesco MA, Tamang S, Yazdany J *et al.* Potential biases in machine learning algorithms using electronic health record data. *JAMA Intern. Med.* 2018;**178**:1544–7. https://doi.org/10.1001/jamainternmed.2018.3763.
84. Mahmood SS, Levy D, Vasan RS *et al.* The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. *The Lancet* 2014;**383**:999–1008. https://doi.org/10.1016/S0140-6736(13)61752-3.
85. Blum A, Wang P, Zenklusen JC. SnapShot: TCGA-analyzed tumors. *Cell* 2018;**173**:530. https://doi.org/10.1016/j.cell.2018.03.059.
86. Parisi GI, Kemker R, Part JL *et al.* Continual Lifelong Learning with Neural Networks: A review. *Neural Netw* 2019;**113**:54–71. https://doi.org/10.1016/j.neunet.2019.01.012.
87. Hoare D, Bussooa A, Neale S *et al.* The future of cardiovascular stents: bioresorbable and integrated biosensor technology. *Adv Sci (Weinh)* 2019;**6**:1900856. https://doi.org/10.1002/advs.201900856.
88. Castelvecchi D. Can we open the black box of AI? *Nature* 2016;**538**:20–3. https://doi.org/10.1038/538020a.
89. Ferrario A, Loi M, Viganò E. Trust does not need to be human: it is possible to trust medical AI. *J Med Ethics* 2020;**47**:437–8. https://doi.org/10.1136/medethics-2020-106922.
90. Tjoa E, Guan C. A survey on explainable artificial intelligence (XAI): towards medical XAI. *IEEE Trans Neural Netw Learning Syst* 2021;**32**:4793–813. https://doi.org/10.1109/TNNLS.2020.3027314.
91. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Med Image Anal* 2022;**79**:102470. https://doi.org/10.1016/j.media.2022.102470.
92. Reddy S. Explainability and artificial intelligence in medicine. *The Lancet Digital Health* 2022;**4**:e214–5. https://doi.org/10.1016/S2589-7500(22)00029-2.
93. Ribeiro MT, Singh S, Guestrin C."why should I trust you?": explaining the predictions of any classifier. 2016. https://doi.org/10.48550/arXiv.1602.04938.
94. Chen H, Lundberg S, Lee S-I. Explaining models by propagating shapley values of local components. 2019. https://doi.org/10.48550/arXiv.1911.11888.
95. Selvaraju RR, Cogswell M, Das A *et al.* Grad-CAM: visual explanations from deep networks via gradient-based localization. *In 2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, p.618–26.
96. Arrieta AB, Díaz-Rodríguez N, Ser JD *et al.* Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. 2019. https://doi.org/10.48550/arXiv.1910.10045.
97. Cruz Rivera S, Liu X, Chan A-W *et al.* Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nat Med* 2020;**26**:1351–63. https://doi.org/10.1038/s41591-020-1037-7.
98. Martin KD, Zimmermann J. Artificial intelligence and its implications for data privacy. *Curr Opin Psychol* 2024;**58**:101829. https://doi.org/10.1016/j.copsyc.2024.101829.
99. Zeb I, Uqaily R, Gonuguntla K *et al.* Multimodality assessment of high- vs. low-gradient aortic stenosis using echocardiography and cardiac CT. *J Cardiovasc Comput Tomogr* 2023;**17**:421–8. https://doi.org/10.1016/j.jcct.2023.09.002.

100. Danad I, Szymonifka J, Twisk JWR *et al.* Diagnostic performance of cardiac imaging methods to diagnose ischaemia-causing coronary artery disease when directly compared with fractional flow reserve as a reference standard: a meta-analysis. *Eur Heart J* 2016;**38**:991–8. https://doi.org/10.1093/eurheartj/ehw095.
101. Achenbach S, Fuchs F, Goncalves A *et al.* Non-invasive imaging as the cornerstone of cardiovascular precision medicine. *European Heart Journal—Cardiovascular Imaging* 2022;**23**:465–75. https://doi.org/10.1093/ehjci/jeab287.
102. Dahl A, Hernandez-Meneses M, Perissinotti A *et al.* Echocardiography and FDG-PET/CT scan in gram-negative bacteremia and cardiovascular infections. *Curr Opin Infect Dis* 2021;**34**:728–36. https://doi.org/10.1097/QCO.0000000000000781.
103. Thackeray JT, Derlin T, Haghikia A *et al.* Molecular imaging of the Chemokine receptor CXCR4 after acute myocardial infarction. *JACC Cardiovasc Imaging* 2015;**8**:1417–26. https://doi.org/10.1016/j.jcmg.2015.09.008.
104. Nagueh SF, Smiseth OA, Appleton CP *et al.* Recommendations for the evaluation of left ventricular diastolic function by Echocardiography: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Eur Heart J Cardiovasc Imaging* 2016;**17**:1321–60. https://doi.org/10.1093/ehjci/jew082.
105. Wagner A, Mahrholdt H, Holly TA *et al.* Contrast-enhanced MRI and routine single photon emission computed tomography (SPECT) perfusion imaging for detection of subendocardial myocardial infarcts: an imaging study. *Lancet* 2003;**361**:374–9. https://doi.org/10.1016/S0140-6736(03)12389-6.
106. Gould KL, Johnson NP. Coronary CT angiography with PET perfusion imaging: hybrid or hype? *JACC Cardiovasc Imaging* 2017;**10**:1371–3. https://doi.org/10.1016/j.jcmg.2016.09.033.
107. Karim R, Blake L-E, Inoue J *et al.* Algorithms for left atrial wall segmentation and thickness—evaluation on an open-source CT and MRI image database. *Med Image Anal* 2018;**50**:36–53. https://doi.org/10.1016/j.media.2018.08.004.
108. Rajiah PS, Reddy P, Baliyan V *et al.* Utility of CT and MRI in tricuspid valve interventions. *Radiographics* 2023;**43**:e220153. https://doi.org/10.1148/rg.220153.
109. Freed BH, Collins JD, François CJ *et al.* MR and CT imaging for the evaluation of pulmonary hypertension. *JACC Cardiovasc Imaging* 2016;**9**:715–32. https://doi.org/10.1016/j.jcmg.2015.12.015.
110. Makowski MR, Rischpler C, Ebersberger U *et al.* Multiparametric PET and MRI of myocardial damage after myocardial infarction: correlation of integrin $\alpha v\beta 3$ expression and myocardial blood flow. *Eur J Nucl Med Mol Imaging* 2021;**48**:1070–80. https://doi.org/10.1007/s00259-020-05034-z.
111. Rischpler C, Nekolla SG, Kunze KP *et al.* PET/MRI of the heart. *Semin Nucl Med* 2015;**45**:234–47. https://doi.org/10.1053/j.semnuclmed.2014.12.004.
112. Stahlschmidt SR, Ulfenborg B, Synnergren J. Multimodal deep learning for biomedical data fusion: a review. *Brief Bioinform* 2022;**23**:bbab569. https://doi.org/10.1093/bib/bbab569.
113. Steyaert S, Pizurica M, Nagaraj D *et al.* Multimodal data fusion for cancer biomarker discovery with deep learning. *Nature Machine Intelligence* 2023;**5**:351–62. https://doi.org/10.1038/s42256-023-00633-5.

Received 17 May 2025; revised 6 July 2025; accepted 8 July 2025. published 17 July 2025