Canadian Journal of Cardiology 40 (2024) 1880–1896

Review

Emerging Analytical Approaches for Personalized Medicine Using Machine Learning In Pediatric and Congenital Heart Disease

Bhargava K. Chinni, MS,[a] and Cedric Manlhiot, PhD[a,b]

[a] The Blalock-Taussig-Thomas Pediatric and Congenital Heart Center, Department of Pediatrics, Johns Hopkins School of Medicine, Johns Hopkins University, Baltimore, Maryland, USA

[b] Research Institute, SickKids Hospital, Department of Pediatrics, University of Toronto, Toronto, Ontario, Canada

See editorial by Avram, et al., pages 1769-1773 of this issue.

ABSTRACT

Precision and personalized medicine, the process by which patient management is tailored to individual circumstances, are now terms that are familiar to cardiologists, despite it still being an emerging field. Although precision medicine relies most often on the underlying biology and pathophysiology of a patient's condition, personalized medicine relies on digital biomarkers generated through algorithms. Given the complexity of the underlying data, these digital biomarkers are most often generated through machine-learning algorithms. There are a number of analytic considerations regarding the creation of digital biomarkers that are discussed in this review, including data preprocessing, time dependency and gating, dimensionality reduction, and novel methods, both in the realm of supervised and unsupervised

RÉSUMÉ

La médecine de précision et la médecine personnalisée, qui consistent à adapter la prise en charge d'un patient à sa situation personnelle, sont des termes connus des cardiologues, même s'il s'agit d'un domaine encore émergent. La médecine de précision repose le plus souvent sur les paramètres biologiques et physiopathologiques qui sous-tendent l'état d'un patient, tandis que la médecine personnalisée utilise des biomarqueurs numériques générés au moyen d'algorithmes. Vu la complexité des données sous-jacentes, ces biomarqueurs numériques sont principalement générés au moyen d'algorithmes d'apprentissage automatique. Nous examinons ici divers facteurs analytiques pris en compte pour générer des biomarqueurs numériques, notamment le prétraitement

The concept of precision medicine is now familiar to most clinicians caring for children with heart disease. Fundamentally, the goal of precision medicine is to target the management of patients to the biological underpinnings of their condition as opposed to managing reasonably similar patients as homogeneous groups with standardized therapy. To date, much of the early focus of precision medicine has been on the use of physiological or genetic biomarkers to target therapy. This approach has been particularly successful in fields such as oncology,[1] but some examples do exist in the field of cardiology, such as the avoidance of clopidogrel in patients who are poor metabolizers of CYP2C19 because of deleterious genetic mutations.[2] In this review, we will explore examples and analytic considerations of personalized medicine, a subtle variation of precision medicine that uses digital biomarkers, generated by algorithms, to individualize the care of children with heart disease. Specifically, we will demonstrate that personalized medicine through digital biomarkers often requires the use of machine learning (ML) and explore the various scenarios in which digital biomarkers can be used clinically. We will review through concrete examples, both classic and emerging ML methods that are used to generate digital biomarkers and highlight some challenges that are unique for children with heart disease. Additionally, we will introduce the emerging concepts of predictive allocation and clinical artificial intelligence (AI) operations, discussing the analytic considerations and challenges associated with the deployment of digital biomarkers for the care of children with heart disease.

From Expert Opinion to Personalized Medicine

Historically, diagnosis, prognosis, and management of human disease relied on the extensive expertise and insights of individual physicians, a practice commonly referred to as expert opinion.[3] Although this approach can offer valuable insights, it has limitations. Subjectivity and potential knowledge gaps could lead to treatment inconsistencies, and reliance solely on individual expertise limits the reliable incorporation

Received for publication May 31, 2024. Accepted July 26, 2024.

Corresponding author: Dr Cedric Manlhiot, Johns Hopkins Hospital, 600 N. Wolfe Street, 1389 Blalock, Baltimore, Maryland 21287, USA. Tel.: +1-410-955-5000.

E-mail: cmanlhi1@jhmi.edu

See page 1893 for disclosure information.

https://doi.org/10.1016/j.cjca.2024.07.026

machine learning. Some of these considerations, such as sample size requirements and measurements of model performance, are particularly challenging in small and heterogeneous populations with rare outcomes such as children with congenital heart disease. Finally, we review analytic considerations for the deployment of digital biomarkers in clinical settings, including the emerging field of clinical artificial intelligence (AI) operations, computational needs for deployment, efforts to increase the explainability of AI, algorithmic drift, and the needs for distributed surveillance and federated learning. We conclude this review by discussing a recent simulation study that shows that, despite these analytic challenges and complications, the use of digital biomarkers in managing clinical care might have substantial benefits regarding individual patient outcomes.

des données, l'évolution et la fragmentation des données (« gating ») en fonction du temps, la réduction de dimensionnalité et les nouvelles méthodes, tant pour l'apprentissage automatique supervisé que non supervisé. Certains de ces facteurs, par exemple les exigences relatives à la taille de l'échantillon et les mesures de la performance du modèle, présentent des difficultés particulières pour les populations petites et hétérogènes de patients atteints de maladies rares, par exemple les enfants atteints de cardiopathie congénitale. Enfin, nous passons en revue les facteurs analytiques utilisés pour le déploiement de biomarqueurs numériques en contexte clinique, par exemple le domaine émergent des opérations cliniques d'intelligence artificielle (IA), les besoins informatiques liés au déploiement, les efforts nécessaires pour améliorer l'explicabilité de l'IA, la dérive algorithmique et les besoins liés aux systèmes distribués de surveillance et de l'apprentissage fédéré. Nous concluons en présentant une récente étude de simulation qui révèle qu'en dépit de ces difficultés et complications analytiques, l'utilisation de biomarqueurs numériques pour la prise en charge clinique promet d'améliorer de façon notable les résultats pour les patients.

of the latest research findings in clinical care.[4] In the latter half of the 20th century, there was a significant shift in medical practice with the emergence of evidence-based medicine (EBM). EBM advocates for making medical decisions based on robust clinical evidence, particularly stemming from randomized controlled trials and meta-analyses to inform treatment.[3] This approach led to significant improvements in patient management and outcomes.[5] However, EBM faced criticism for potentially disregarding individual variations in patients and the heterogeneity in patient response to treatment. On average, treatment recommendations estimated from extensive studies based on population-level evidence might not perfectly or consistently apply to every individual patient, as it is not considering factors that could influence treatment outcomes.[4] In some cases, one or more risk factors have enough impact on treatment response and outcomes to warrant stratified management of patients, but even in these cases, significant heterogeneity in treatment outcomes can remain. This is when the concept of precision and personalized medicine emerged, with advancements in genetics and analytic methods. Both approaches aim to provide precise treatment plans for patients enhancing their diagnosis, prognosis, and treatment choices.[4-7]

Emergence of Precision and Personalized Medicine

Precision medicine requires the profiling of patients using their clinical, physiological, and omics data (biological dimensions) such as genomics, metabolomics, proteomics, and transcriptomics to optimize medical treatment while minimizing any adverse events.[4,7] Personalized medicine extends this approach by using algorithms to generate personalized insights by integrating information from complex biological dimensions in combination with clinical, psychological, environmental, and socioeconomic dimensions and identifying hidden patterns to personalize patient management and further enhance treatment response and outcomes.[4-7] These insights, when generated by algorithms, are sometimes referred to as "digital biomarkers". Although these terms are often used interchangeably, we use precision medicine as targeting treatment based on patients' biological characteristics, whereas personalized medicine uses a broader array of data to tailor individualized care.

Why Does Personalized Medicine Often Require AI?

The biggest challenge with precision medicine is the need to integrate and use the large swath of information and data elements that might affect a disease prognosis or outcome for a given patient. Previous studies of the decision-making process of medical professionals have shown that decision making is done by considering key[5-7] facts about patients' situations in front of them and that their medical training and experience, sometimes referred to as intuition or *gestalt*, rapidly integrate the rest of the information available to arrive at a decision.[8] Traditional approaches to create digital biomarkers have also used a similar approach, focusing on a few key factors to generate potentially useful outputs. This is true of regression models, which are usually fitted on a small number of features, point systems, clinical decision trees, and so forth. The problem is that this approach invariable reduces the complexity of clinical scenarios and can leave crucial information away from the clinical decision-making process. In a previous study, we have made such a demonstration in the context of cardiopulmonary exercise testing for the prognostication of patients with heart failure (HF). The current practice with cardiopulmonary exercise tests is to take staged and breath-by-breath data and consolidate it into a small number of summary indices for clinical interpretation. We have shown that prediction models for 1-year outcomes using these indices were much less accurate than prediction models generated using the staged and breath-by-breath data directly.[9] The challenge here is that almost invariably, traditional statistical algorithms such as regression models will require the oversimplification of clinical scenarios, given the limits in the type and quantity of inputs that can be accommodated. In consequence, the digital biomarkers that are generated through these methods are likely to have suboptimal

performance and limited clinical utility. In these cases, AI will be needed.

Role of AI and ML

AI is and will remain a transformative force in health care, revolutionizing how we diagnose, treat, and manage diseases. In the realm of personalized medicine, AI plays a pivotal role in deciphering the complexities of individual patient data to deliver tailored treatment solutions.[4,6] At its core, AI encompasses a diverse set of technologies and methodologies that enable machines to perform tasks that typically require human intelligence.[6] In medicine, most AI contributions fall under the umbrella of ML. There are a few characteristics of ML algorithms that make them highly suitable for personalized medicine. First, ML algorithms can be trained on vast amounts of data, far above the amount that can be considered by traditional statistical methods. ML algorithms can also consider data from novel data sources ranging from genetic profiles, imaging, waveforms, electronic health records (EHR), and physiological data from wearable devices.[4,6,10-14] These algorithms can also detect subtle, not clinically evident, abnormalities and provide clinicians with actionable insights to guide treatment decisions. Furthermore, ML facilitates the integration of nonclinical data sources including lifestyle factors, environmental exposures, and geospatial factors.[4,7] Beyond the ability to consider nontraditional data inputs,[15] the other key strengths of ML in this context lies in the ability to uncover nonclinically evident patterns in data that may elude human perception; the ability to generate higher order features through complex interactions in data; and, finally, the ability to incorporate variables with very abnormal distributions.[14] In many situations, ML will be necessary to generate high-quality digital biomarkers to guide treatment in a personalized approach.

ML Methods for Digital Biomarkers

Classes of ML algorithms

ML comprises a suite of algorithms that use sample data to identify patterns that can subsequently be applied for predictions or classification. There are multiple classes of ML algorithms, each with their own structure and advantages. An in-depth review of the various classes of ML algorithms is beyond the scope of this review, but multiple excellent reviews exist on this topic. ML algorithms can be classified broadly into boundary-based algorithms (support vector machine, Bayes),[16] tree-based algorithms such as random forest,[17] and its multiple permutations (extreme gradient boosting [XGB],[18] light gradient boosting,[19] CatBoost,[20]) and artificial neural networks (ANN).[21]

Optimizing ML Models

ML methods provide a diverse array of model parameters, facilitating hyperparameter tuning[22] for optimizing model performance and enhancing predictive accuracy. These hyperparameters encompass various aspects of the model, such as the learning rate, regularization parameters, tree depth, number of tree estimators, feature and observations subsampling rate, choice of kernel, number of hidden layers in ANN, and the number of neurons in each layer. Tree ensemble techniques such as boosting, bagging, and bootstrapping fortify models for greater robustness and generalizability, thereby bolstering predictive power. Notably, employing shallow trees with reduced depth, as seen in methods like XGB, can minimize errors effectively, leading to superior performance outcomes. By leveraging these advanced techniques alongside hyperparameter tuning, ML methods often surpass linear classification methods including logistic, lasso, and ridge regressions by accommodating nonlinear relationships, feature interactions, complex data distributions, mitigating overfitting, and ultimately achieving higher predictive accuracy.

Managing Overfitting

Although there are multiple obvious advantages to using ML when it comes to individualized medicine, there are 2 important drawbacks that should be addressed regarding ML. First, is the risk of overfitting, a condition in which ML algorithms overtrain. Given that ML algorithms can handle a very large degree of complexity in the feature space, model overtraining can lead to overfitting, a phenomenon in which the model "learns" the feature space as opposed to the underlying rules that govern the association between features and outcomes. Overfitted prediction models are thus unable to generalize to unseen data. There are several methods to prevent overfitting of ML models that can otherwise lead to underfitting where the algorithm fails to fully learn the associations between the feature space and outcomes. Regardless, the fitting problem (also known as the bias-variance tradeoff) is an important consideration when choosing ML for the creation of digital biomarkers.

Explainability and Interpretability

The second caveat about using ML in this context is the issues of explainability and interpretability.[23] Not all ML algorithms are explainable; this means that the direct relationship between the inputs and outputs of an algorithm cannot be described. Generally, the more complex an algorithm, the better the performance but lower the explainability. There are many emerging statistical methods to understand the algorithm decision-making process (eg, variable importance, SHAP [**Sh**apley **A**dditive Ex**p**lanations] values), which improves their interpretability (sometimes referred to pseudoexplainability); however, it is important to remember that interpretability/pseudoexplainability and explainability of ML models are not the same and cannot be used interchangeably, particularly in the context of regulatory approval of digital biomarkers.

When to Use Logistic Regression for Digital Biomarkers

Because of these 2 important limitations, in some situations, logistic regression (LR) will be favoured for the creation of digital biomarkers. LR is a powerful supervised classification method valued for its simplicity, computational efficiency, and adaptability to modest datasets.[24,25] LR derives interpretable results through coefficients that explain the

relationship between features and the target variable. LR segments the patients into groups under the assumption of a linear relationship between features and the log-odds of the target variable, and the features are independent of each other.[26] Given its simplicity and full explainability, LR should be favoured for situations in which the inputs are limited in type and numbers (particularly binary and categorical data elements) and when explainability is mandatory.

The Different Roles of Digital Biomarkers in Personalized Medicine

Digital biomarkers can be used in many ways to achieve the goals of personalized medicine, and they can be designed to operate across all 4 levels of involvement of AI systems in medicine: namely, automation, assisting, augmenting, and autonomous systems.[14] To illustrate each level of involvement in personalized medicine, we will explore examples pertinent to addressing challenges in cardiovascular disease (CVD) (Table 1). Systems can be categorized based on 2 axes: the degree of complexity of their tasks and the degree of adaptability needed. Automatic and assisting systems typically have low complexity and require limited adaptability, whereas augmenting and autonomous systems handle more complex tasks with greater adaptability needs, making them crucial for the implementation of personalized medicine.[14]

Use Cases from Cardiology

In cardiology, many AI systems use digital biomarkers to perform automatic tasks. For example, ML algorithms can analyze electrocardiography (ECG) to detect abnormal heart rhythms[27,28] or to screen patients for conditions such as coronary artery disease (CAD), myocardial infarction (MI), and congestive heart failure (CHF).[29] The system leverages convolutional neural networks (CNNs) and Gabor-CNN models to diagnose each of these conditions with 98.5% overall accuracy. Assisting AI systems function by analyzing input data to generate relevant insights on a given patient, this can include predicting specific diagnosis, risk of complications, likely response to treatment, forecast of disease progression.[4-7,10-13] These insights can be used to generate digital biomarkers that allow for patient prioritization, greater frequency and intensity of monitoring, early intervention, and personalized preventive measures.[10,30,31] Integrated with electronic medical records in primary care, **Predict**ion of **C**ardio**v**ascular **D**isease (PREDICT-CVD) aimed to improve documentation of CVD risk factors and increase the intensity of testing and monitoring for high-risk patients. The study design involved auditing patient records before and after system implementation, demonstrating a significant increase in documented CVD risk after using PREDICT-CVD.[32] Digital biomarkers can also be used in augmenting systems, which go beyond simple recommendations and provide insights and data that might not be otherwise available to the end user. For example, a deep-learning algorithm trained to provide plaque characteristics from coronary computed tomography (CT) angiography as opposed to solely identifying blockages.[33] This is crucial information, as plaque volume and composition are crucial indicators of heightened risk for future heart attacks.[34-37] Although still an emerging field, digital

Table 1. Levels of AI involvement in cardiology*

AI involvement level	Description	Adaptability	Complexity	Example in cardiology	Benefits	Challenges
Automated	Performs tasks with minimal human intervention	Low	Low	ECG analysis for arrhythmia detection	Improved efficiency, reduced errors	Limited decision-making capabilities
Assisting	Provides recommendations based on patient data analysis	Low	Medium	Medication and intervention suggestions	Improved decision making, optimized patient flow	Requires human overview
Augmenting	Offers insights beyond basic recommendations	High	Medium	Plaque characterization in coronary arteries	Early identification of high-risk patients	Requires human expertise for interpretation
Autonomous	Performs complex tasks with minimal oversight	High	High	Chest X-ray analysis for cardiovascular disease diagnosis	Faster diagnosis, improved resource allocation	Requires robust safety measures and ongoing monitoring

AI, artificial intelligence; ECG, electrocardiogram.
*This table summarizes the different levels of involvement that AI systems can have in cardiology. Each level is characterized by its adaptability, complexity, a relevant example of its application in diagnosing or managing cardiovascular disease, the benefits it offers health care professionals, and the challenges associated with its implementation.

biomarkers will also eventually make their way in autonomous systems; those are systems that can perform an action without the intervention of a human operator. Examples of autonomous system using digital biomarkers will include AI-powered automated implantable cardioverter defibrillators.

At this point we have demonstrated that the transition from EBM to personalized medicine requires the use of digital biomarkers. In many cases, ML algorithms will be needed to generate those biomarkers, although because of some limitations of ML algorithms, there is a narrow group of cases that would benefit from a simpler approach such as LR. We have then demonstrated that those digital biomarkers can be used as part of a personalized medicine approach in 4 distinct types of AI systems. The next section will focus on specific examples from the field of pediatric heart disease and discuss some analytic considerations regarding the development to digital biomarkers.

Translating Clinical Questions Into AI-Friendly Framework for Personalized Medicine

The first important analytical challenge with ML is to translate clinical problems and questions into a framework that can be used by existing ML algorithms (Fig. 1). ML uses such techniques as supervised learning (classifying diseases), unsupervised learning (finding patterns), and reinforcement learning that are pivotal in tasks ranging from diagnosis to treatment prediction and patient monitoring. Key steps in building AI models involve defining objectives, gathering and preprocessing data, and training and validating ML models with large datasets. The choice of AI methods heavily depends on the type of data used. For instance, cardiology benefits from DL for analyzing images using CNNs and signals using recurrent neural networks (RNNs). The integration of unsupervised learning and multimodal data fusion holds promise for identifying subphenotypes, facilitating tailored treatments in personalized medicine. Fundamentally, to translate a clinical problem into in a ML framework, one must first define a feature space, and in supervised or reinforcement learning problems further define an outcome and a horizon (the horizon is the length of time between the acquisition of the data included in the feature space and the outcome). Understanding these core concepts is essential for using ML for the goal of personalized medicine. A previous review[38] dives into the AI workflow, focusing on key ML algorithms and how they measure success in cardiovascular care. Additional considerations such as feature space, time dependency and gating mechanisms, and data dimensionality are necessary to translate complex clinical problems into AI treatable challenges effectively.

Data Processing and Feature Selection for Personalized Medicine

Clinical data, especially omics data, frequently reside within high-dimensional feature spaces in which each dimension captures a distinct patient characteristic.[39] In health care, clinical data originates from such diverse sources as ECG recordings and echocardiograms, capturing diverse types of information in image format and transformed into numerical format. These datasets commonly encompass a mixture of continuous variables (blood pressure [BP] readings), categorical variables (medical diagnosis), and binary variables (presence or absence of a specific comorbidity), reflecting the multifaceted nature of patient information in clinical settings. In such high-dimensional spaces, the task for ML models is to identify the pertinent features essential for making precise predictions to estimate the risk. This involves using data-mining methods, exploring through the vast array of dimensions to implement feature selection and pinpoint those that hold the most significance in influencing outcomes accurately. By effectively identifying and selecting relevant features, ML models can enhance their predictive power, thereby contributing to more informed clinical decision making and personalized clinical decision making.

Time Dependency and Gating

The investigation of time dependency and gating becomes paramount when analyzing diseases characterized by evolving patterns over time. This necessitates the use of specialized ML techniques, such as RNN splitting the data into specific time intervals (gating) capable of effectively learning patterns from sequential data.[40] RNNs are particularly adept at capturing temporal dependencies[41,42] inherent in the progression of diseases, enabling more accurate and insightful analyses of dynamic medical processes. AI-driven algorithms analyzing gated ECG signals[42,43] can enhance the diagnosis of CVDs such as pulmonary hypertension,[44] HF,[45,46] atrial fibrillation,[47] and hypertrophic cardiomyopathy (HCM).[48] AI models tasked with diagnosing CVD can use time gating to examine patterns in ECG,[41] physiological features such as heart rate and BP levels. This approach not only facilitates accurate diagnosis but also enables the prediction of future risks by tracking the evolution of these features over time. Through the integration of time gating, AI emerges as a valuable tool in addressing complex clinical challenges; moreover, these tools can be customized for integration of wearable devices, enhancing their utility in real-world health care settings.

Data Fusion and Dimensionality Reduction

Principal component analysis

Personalized medicine in cardiology relies on a multitude of heterogeneous data dimensions—including imaging, biology, waveform, psychology, environment, and socioeconomics—to understand individual health and tailor treatments. Often, the number of dimensions measured is much higher than the number of patients available to train an algorithm. This phenomenon often leads to the injection of substantial noise in prediction models and tend to result in poor model performance. To overcome this, techniques such as principal component analysis (PCA) were employed to reduce complexity and focus on the most relevant data. Linear techniques such as PCA project patients onto a lower-dimensional space while retaining the most significant variance. This simplifies visualization and helps identify clusters or trends associated with specific diseases. However, PCA assumes linear relationships among features, which might not always hold true in biological systems.

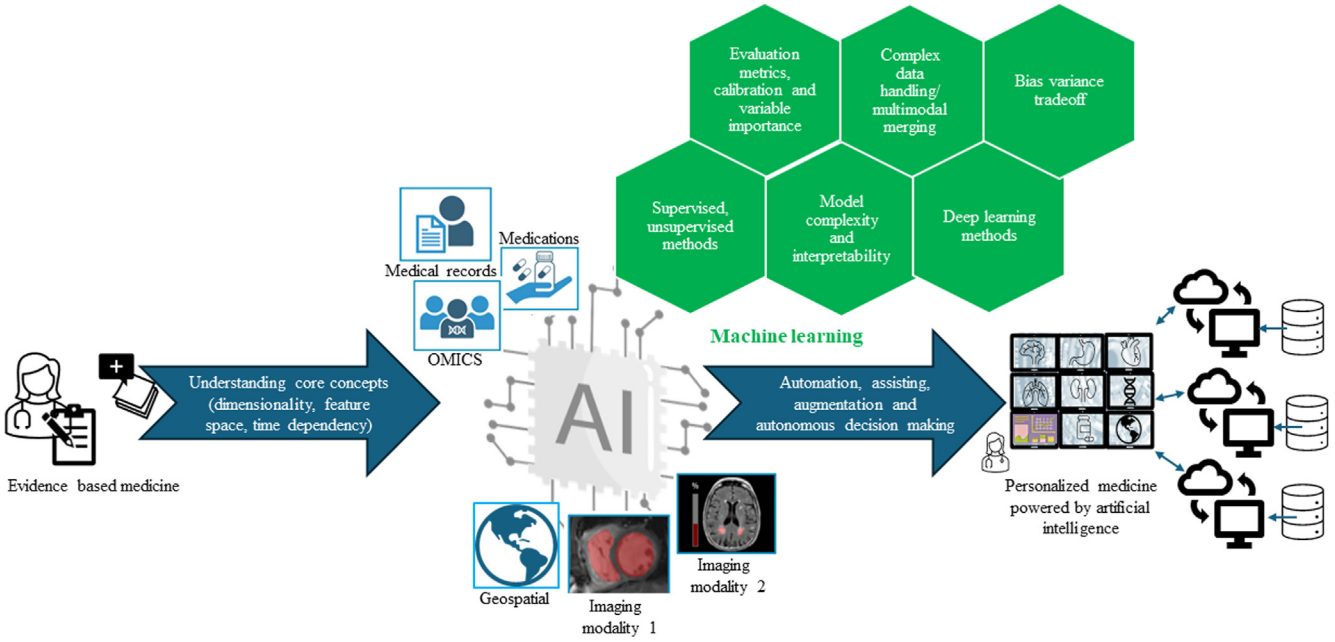

Figure 1. The journey of AI in medicine. This figure illustrates the evolution of AI in health care, from evidence-based medicine to personalized medicine powered by AI. The **arrows** highlight the importance of understanding core AI concepts (**left arrow**) such as dimensionality, feature space, and time dependency and the range of AI applications (**right arrow**) from automation to autonomous decision making. It also highlights the key strengths, methods, and analytic challenges associated with implementation of achieving personalized medicine. The promising horizons of personalized medicine encompass advancements such as clinical AI operations, federated learning, and distributed surveillance. These innovations pave the way for a global model that transcends data security concerns. AI, artificial intelligence.

Nonlinear Techniques

For capturing nonlinear relationships, techniques such as t-distributed Stochastic Neighbour Embedding (t-SNE) and Uniform Manifold Approximation and Projection (UMAP) are valuable tools for grouping similar patients in the lower dimensions and aiding in cluster analysis.[38,49,50] These methods focus on preserving the local and global structure of the data, allowing for visualization of complex, nonlinear relationships among features that might be crucial for understanding disease mechanisms.

Feature selection with Boruta dimensionality reduction can also enhance model performance through the removal of redundant features, thereby mitigating the curse of dimensionality and improving generalization. In addition, a reduced feature set enhances interpretability, facilitating a deeper understanding of the relationships between features and the outcome. Boruta, a wrapper method, leverages random forest (RF) algorithm to identify key features by comparing them with randomly shuffled versions (shadow features).[51] Features surpassing noise levels are retained, effectively reducing dimensionality while focusing on the critical features with substantial predictive capacity. This capability renders Boruta particularly valuable in pinpointing critical biomarkers, as the retained features are likely pivotal in elucidating the relationship between omics variables and the CVD outcome.

Novel AI Methods Using Unsupervised ML

Unsupervised ML is a data-driven approach that examines unlabelled datasets.[52,53] It is used for various purposes such as extracting generative features, identifying meaningful trends and structures, grouping results, and exploration. Common tasks in unsupervised learning include clustering, density estimation, feature learning, dimensionality reduction, finding association rules, and detection of anomalies.[52]

Clustering algorithms like PCA, K-means, hierarchical clustering (HC), and UMAP[53-55] can identify patient subgroups with similar characteristics, aiding in risk stratification. Murray et al.[55] investigated the application of UMAP to identify subgroups of patients with distinct vascular characteristics in early hypertension. The analysis revealed 3 distinct patient subgroups: termed arterially stiffened, vasoprotected, and nondippers. The arterially stiffened group exhibited the highest BP (150/100 mm Hg) and minimal nocturnal dipping, whereas the vasoprotected group had the lowest BP (120/80 mm Hg) and the greatest nocturnal dipping. The nondippers displayed intermediate BP (130/90 mm Hg) with no nocturnal dipping and elevated BP variability. These findings suggest the potential of UMAP to identify patient subgroups with early hypertension based on underlying vascular differences, potentially informing personalized treatment approaches.

Clustering Analysis in Heterogeneous Diseases

Clustering analysis, a powerful data mining method and an unsupervised ML technique, is primarily used to groups observations sharing similar characteristics. This is particularly helpful in heterogeneous diseases such as dilated cardiomyopathy (DCM), in which patients with varying underlying causes usually receive similar treatments. In a previous study,[56] UMAP and PCA were used to analyze a feature rich

dataset including patients' genotype, phenotype, and the cardiac transcriptome. Clustering algorithms such as K-means, (HC)[54] and Gaussian mixture models[57] were then applied to this reduced-dimensionality data, to identify distinct patient subgroups representing different underlying pathophysiology. These subgroups differed in transcriptomics, cardiac metabolism, cardiomyocyte function, and inflammatory signature. This approach, using dimensionality reduction and clustering, has the potential to pave the way for personalized medicine by identifying unique disease pathways in distinct patient subgroups and their associate biomarkers, leading to more targeted therapies.

Novel AI Methods Using Supervised ML

As previously noted, there are multiple types of supervised ML algorithms. Algorithm selection considers factors such as maximizing evaluation metrics, interpretability, model size, and computational complexity. Ensemble models and tree-based models are favoured for interpretability, whereas considerations for model size and computational complexity are vital for implementation in wearable devices.[58] This section delves into the latest advancements customized for various data modalities encountered in cardiology, ranging from imaging to wearable technology and beyond. By exploring these innovative methods, we uncover how ML is reshaping the landscape of personalized cardiovascular medicine, propelling the field toward more accurate and effective patient care.

Cardiac imaging data such as echocardiograms, magnetic resonance imaging (MRI), CT, and coronary angiograms rich in visual information, are prime targets for DL techniques.[59] CNN has emerged as the dominant force in this image analysis domain, holds immense potential for automated analysis. Bui et al.[60] focused on developing an AI system called DeepHeart CT for accurate and rapid segmentation of cardiovascular structures in cardiac CT angiography images using DL. They trained ML using a large clinical dataset with computer-generated labels and employed a reverse ranking strategy to assess segmentation quality. The proposed framework achieved high segmentation accuracy across various structures including left and right ventricles, atria, and myocardium, with metrics such as a high median Dice score of 0.90, low median Hausdorff distance of 7 mm, and low mean surface distance of 0.80 mm, demonstrating its potential for large-scale medical imaging applications. In contrast, Alkhodari et al.[61] investigated the efficiency of DL models in identifying valvular heart diseases (VHDs) through phonocardiography recordings. They designed an ANN combining CNN and RNN based on bidirectional long short-term memory (LSTM) and achieved high performance in identification of VHD, with metrics such as an overall Cohen's kappa of 97.9%, accuracy of 99.3%, sensitivity of 98.3%, and specificity of 99.6%, thus demonstrating the promise of DL models for early diagnosis and prevention of cardiac abnormalities.

Network Analysis in Cardiovascular Diseases

Cardiovascular diseases often are a complex interplay among genes, metabolites, and proteins and untangling this web of interactions is crucial for developing new precision treatments. Recent studies have used advanced network analysis techniques to shed light on these intricate biological processes. Unlike traditional dimensionality reduction methods that focus on importance of features, network analysis considers the interactions among various biological entities, allowing for a comprehensive view of the system. This approach highlights 3 main network types: protein-protein interaction networks for drug target discovery, expression-based networks for identifying co-regulated genes in disease, and gene regulatory networks for understanding disrupted cellular processes.[62] Graph-theory concepts are used to represent interactions, as nodes and edges. By analyzing the structural characteristics of these networks, such as node degree and clustering coefficients, researchers can gain insights into the underlying biology of diseases[63,64] and create personalized management strategies. Leveraging publicly available gene expression (RNA-seq) and microRNA data, Pepe et al.[65] performed a meta-analysis across 6 studies encompassing patients with HF, ischemic cardiomyopathy (ICM), DCM, and HCM, alongside healthy controls. A partial correlation coefficient (PCC)-based algorithm was then used to construct disease specific gene regulatory networks by identifying coexpressed genes potentially regulated by the same transcription factors (TFs). By integrating results across all datasets using a meta-analysis approach, the study revealed key regulatory TFs and pathways potentially involved in development of disease, offering valuable insights for potential therapeutic targets. Multimodal learning frameworks are being developed to combine clinical, imaging, and genomics: data for a more comprehensive understanding of CVD.[66-69]

RNN in Physiological Signal Analysis

RNN models are particularly well suited for analyzing physiological signals[70] including ECG, electromyogram (EMG), electroencephalogram (EEG), and electrooculogram (EOG), such as analyzing nonlinear temporal sequences of clinical events because of their ability to capture complex temporal dependencies within data. Choi et al.[71] investigated the application of RNN models in predicting the early onset of HF by converting clinical event sequences and related time-stamped data into pathways relevant to disease detection. Using longitudinal EHR data, which encompasses various clinical encounters and treatments, poses challenges because of their complexity and irregular sampling frequency. Unlike conventional methods that rely on aggregate features, RNN models capture temporal relations among disaggregated features, offering a promising approach for early disease detection. These studies[71,72] compare the performance of RNN models with traditional ML approaches and emphasize the significance of early detection of HF,[8] detection of arrhythmias from ECG data,[72] considering its high morbidity, mortality, and health care costs. By leveraging DL techniques, particularly RNN structures such as LSTM and gated recurrent units (GRUs), researchers can enhance our understanding of progression of disease and prediction of prognosis.

RNN Application in Wearable Devices

RNN application to wearable devices generating continuous physiological features (eg, heart rate, BP, and respiration rate) for handling time series information has emerged as a

promising avenue for advancing health care-monitoring capabilities and aiding in diagnosis of disease, especially in elderly care,[58,73-76] allowing for personalized medicine and post-treatment monitoring. With the global increase in the elderly population and the rising importance of chronic and acute diseases, the medical industry is witnessing a significant shift toward point-of-care diagnosis and real-time health monitoring. Despite challenges, such as motion artifacts, advancements in electronic and nanomaterials have led to the development of implantable devices, further enhancing diagnostic and prognostic capabilities.[73] Hughes et al.[76] evaluated the effectiveness of RNNs in detecting arrhythmias; the sensitivity of the classifier was found to be 92%, and specificity was measured at 95%. These high-sensitivity and specificity values demonstrate the reliability of RNN-based classifiers in distinguishing between normal and abnormal cardiovascular activities.

Special Analytical Considerations for Pediatric and Congenital Heart Disease

Small sample sizes and heterogeneous patient populations have been a consistent challenge in pediatric and congenital heart disease research, and those challenges persist in the training of models for the development of digital biomarkers. Sample size considerations are different for ML algorithms than they are with traditional statistical methods, but they are no less important. Generally, for ML algorithms there is a direct and exponential association between the size of the training sample and the achievable performance of the algorithm. This association is partially because of the increased diversity of cases and controls that are used in the training dataset and to the increase degree of algorithmic complexity afforded by larger samples. Smaller training cohorts will generally result in more limited performance, and large cohorts are generally necessary to achieve very high degree of performance. Larger training cohorts are also needed for many algorithms developed using inputs with highly complex structure (eg, images). It is important to note that the effective sample size may be significantly smaller than the total cohort size, particularly when dealing with rare disease events.[77] For instance, in studies with a low outcome prevalence, such as 0.1%, even large sample sizes may yield a small effective sample size because of the rarity of outcome events.[77] Larger sample sizes offer numerous advantages, allowing for robust selection of features, more complex ML models, and exploration of a wider range of research questions. With limited sample sizes, stronger modelling assumptions must be made, which may affect the reliability and generalizability of the findings.

Fundamental Metrics and Techniques to Balance Model Bias and Variance

Although the area under the curve (AUC) serves as a fundamental metric for algorithm performance, it has its limitations in the context of personalized medicine, particularly when dealing with rare events, which are common in pediatric cardiology. With rare events, a few misclassified observations would have a very severe effect on the AUC. To address these challenges, researchers[78,79] have focused on balancing model bias and variance, employing techniques such as regularization, cross validation, and ensemble learning. A model with high bias may miss crucial patterns in patient data, leading to inaccurate predictions. Conversely, high variance can result in an overly complex model that performs poorly on unseen data.

Advanced Evaluation Metrics

Advanced evaluation metrics such as area under the receiver operating characteristics curve (AUC-ROC), precision area under the curve (PR-AUC) and calibration metrics provide a more comprehensive assessment of aggregated model performance,[80-82] considering such factors as class imbalance and discriminative ability. Calibration metrics such as calibration slope and calibration-in-the-large are crucial to consider, particularly for rare events, and is often more informative than AUC. The calibration slope, ideally equal to 1, reflects how well a model's predicted probabilities match observed outcomes in the dataset.[77] Deviations from 1 indicate a need for recalibration, which involves adjusting the model to improve its generalizability to new data. Calibration-in-the-large assesses the overall difference between predicted and observed outcomes. Finally, research emphasizes the importance of validating models on external validation beyond the internal validation[83,84] to identify overfitting. Overfitting can be measured by the difference in performance between the training and validation cohorts. This ensures the model's generalizability to new patients and avoids overfitting to the training data. In many instances, however, external validation cohorts are not available (particularly for rare diseases); in those cases, computational methods can be used for model validation, although those provide an inferior validation standard and should be avoided when possible.

Table 2 provides a comprehensive overview of the application of various AI methods in cardiology, detailing how these methods address key clinical problems such as disease detection, risk stratification, arrhythmia detection, patient subtype discovery, and long-term patient monitoring. It highlights the specific AI algorithms used and explains the types of input data used, outcome variables, performance metrics, and the applications into personalized treatment strategies. In addition, Table 2 underscores the time horizon for each application, ranging from short-term diagnostic tools to long-term risk prediction models, and identifies key challenges faced in each domain. This structured approach facilitates a clearer understanding of how AI can enhance cardiology practices and improve patient outcomes.

Clinical AI Operations: Analytical Considerations for Digital Biomarkers Postdeployment

The fascinating journey of AI transforming health care is not without its challenges and opportunities. A SWOT (strength, weaknesses, opportunities, threats) analysis (Fig. 2) reveals the strengths of AI in improving diagnostics, personalizing treatment plans, and streamlining drug development. However, weaknesses such as reliance on big data, potential for bias, and ethical considerations require careful attention. It is now increasingly being recognized that to be effective and useful, digital biomarkers must be deployed within an appropriate

Table 2. AI methods in personalized medicine for cardiology*

Clinical problem	AI method selection	Input data	Outcome variable	Performance metrics	Applications in personalized medicine	Time horizon	Specific AI tools/ algorithms	Key challenges
Disease detection and risk stratification	Supervised learning (classification, regression)	Patient demographics (age, sex, etc), Medical history (comorbidities), laboratory results (blood tests, biomarkers), imaging data (echocardiograms, MRI scans)	Disease presence, cardiovascular events	Accuracy, precision, recall, AUC-ROC	Automate disease detection from imaging data, improving diagnostic accuracy; develop personalized risk scores for cardiovascular events, informing targeted preventive strategies	Short-term (diagnosis) to long-term (risk prediction)	Random forest, SVM, logistic regression, XGBoost, CatBoost	Data heterogeneity, limited labelled data
Arrhythmia detection and classification	Supervised learning (classification)	ECG signals	Arrhythmia type	Accuracy, sensitivity, specificity	Enable real-time arrhythmia detection for personalized monitoring and treatment decisions; improve early detection of arrhythmias, allowing for prompt interventions and potentially reducing complications	Short-term (diagnosis)	CNN, XGBoost, CatBoost	Signal noise, real-time processing requirements
Patient subtype discovery	Unsupervised learning (clustering)	Clinical data (symptoms, medications), biomarker data (gene expression), imaging data (structural features)	Patient subtypes	Silhouette score (clustering quality)	Identify patient subgroups with distinct disease presentations or responses to therapy; guide development of targeted treatment approaches based on specific patient subtypes	Mid-term to long-term	K-means clustering, Gaussian mixture models, hierarchical clustering, UMAP, t-SNE	Defining optimal number of clusters, feature engineering, high-dimensional data
Long-term patient monitoring and early intervention	Supervised learning (time-series analysis)	Real-time physiological data (wearables), EHR data, patient behavioUr data (activity level)	HF decompensation	AUCROC, AUPRC	Continuously monitor patients at risk for HF decompensation; predict potential deterioration and enable early interventions to prevent complications; personalize patient monitoring plans based on individual risk profiles	Short-term (diagnosis) to long-term (risk prediction)	RNN, LSTM, GBM, XGBoost, CatBoost	Data privacy, long-term data integrity and storage

AI, artificial intelligence; AUC-ROC, area under the curve of the receiver operating characteristic; AUPRC, area under the precision recall curve; CNN, convolutional neural networks; ECG, electrocardiogram; GBM, gradient boosting machines EHR, electronic health record; HF, heart failure; LSTM, long-/short-term memory; MRI, magnetic resonance imaging; RNN, recurrent neural networks; SVM, support vector machine; t-SNE, t-distributed stochastic neighbour embedding; UMAP, uniform manifold approximation and projection.

*This table outlines the application of various AI methods in addressing key clinical problems within cardiology. Each row specifies a clinical problem, the AI method used, types of input data, performance metrics for evaluation, specific applications within personalized medicine to enhance patient outcomes, the time horizon for achieving these applications, specific AI tools or algorithms used, and key challenges faced.

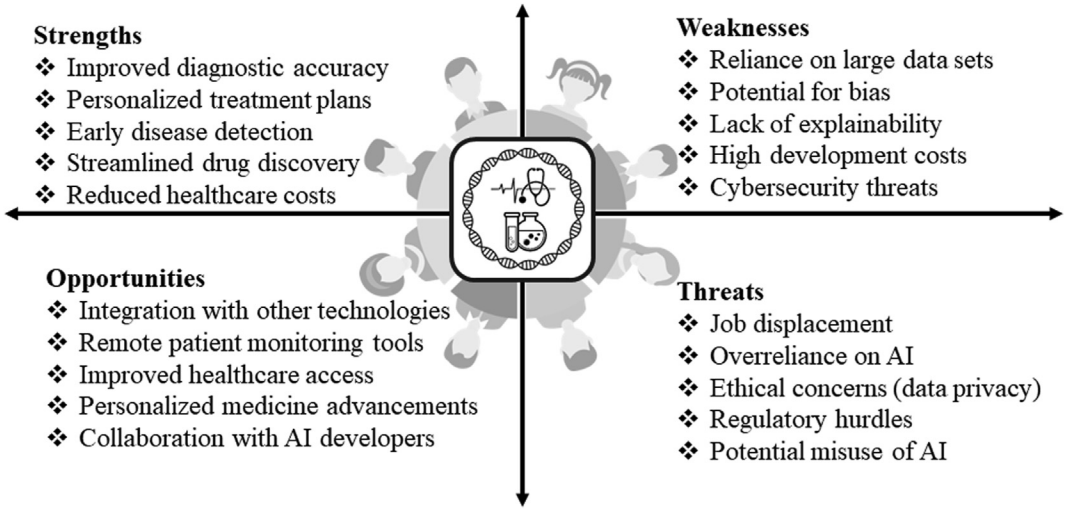

Figure 2. SWOT analysis of artificial intelligence in health care. This SWOT analysis categorizes the strengths, weaknesses, opportunities, and threats associated with the application of AI in health care. The analysis highlights the potential of AI to improve diagnostics, personalize treatment plans, and streamline drug development, while acknowledging challenges such as reliance on big data, potential for bias, and ethical considerations. AI, artificial intelligence; SWOT, strengths, weaknesses, opportunities, threats.

ecosystem to maximize the likelihood of successful clinical implementation. The development and study of this ecosystem is sometimes termed "clinical AI operations" (Table 3).

Computing Resources for Deployment Environment

For successful clinical integration, AI models need seamless integration into existing workflows, with appropriate resources. Clinical AI-Ops encompasses such tasks as model deployment, monitoring, and version control, ensuring smooth and reliable use of AI tools by health care professionals.[85] Integration with EHRs and clinical decision support systems (CDSS) is crucial for facilitating real-time access to AI-powered insights during patient care.[86] The proliferation of cloud computing infrastructure and advanced programming tools has democratized development and deployment of AI.[87]

Cloud computing facilitates molecular modelling, omics data analytics, and integration and interpretation of phenotypic data. Integration of big data and cloud computing technologies is necessary to enable scalable analytics infrastructures such as the Hadoop Distributed File System (HDFS) for storage and Spark Machine Learning Libraries (MLlib) for analysis.[87] Cloud-based platforms, such as the National Cancer Institute Cloud Pilots program FireCloud enable secure and reproducible analysis of shared datasets, reducing the reliance on costly local infrastructure and facilitating effective data management. Using these advanced pipelines and workflows on cloud platforms, researchers can leverage cost-effective computing resources and streamline data analysis processes.

Explainable AI and Reporting Feature Importance

In addition to meticulous ML methodology selection, successful implementation of ML techniques in cardiology, as well as other domains, demands thoughtful consideration of various factors to ensure their efficacy and reliability in clinical settings. Although traditional risk factors identified through clinical studies lay a robust foundation for cardiovascular risk prediction, ML models often unveil additional variables that contribute to their performance. However, the clinical relevance of these variables must be scrutinized despite their statistical significance or prediction importance within the model.

Techniques such as feature permutation using Boruta[51] and SHAP[88] offer invaluable insights into each feature's contribution to ML predictions, facilitating interpretation and validation. Evaluation of ML performance extends beyond basic metrics like accuracy. Calibration becomes pivotal to ensure the model's predicted risk aligns accurately with the actual risk of an event.[89,90] Approaches such as Platt scaling and isotonic regression[89] can refine predicted probabilities, thereby bolstering the reliability of risk assessments.[89,90]

Bias Mitigation

ML models are susceptible to inheriting biases from their training data, potentially hindering generalization to new populations. Training bias can arise from factors such as the data used to train the ML and the way the data are preprocessed.[91] For example, if the training data include more data from one group of people than another, the algorithm may be biased toward that group. Mitigation strategies using diverse patient populations in combination with strategies such as regularization, cross validation, ensemble learning, data augmentation, and the incorporation of multicentre datasets during training are instrumental.

Explainable AI

Enhancing the explainability of ML models is imperative for fostering trust among clinicians and patients in cardiology practice. Explainable AI (XAI)[92,93] methods such as local interpretable model-agnostic explanations (LIME)[93,94] and

Table 3. AI advancements in cardiology

Advancement	Description	Benefits	Challenges
Clinical AI operations	Ensures smooth and reliable use of AI tools by health care professionals	Integrates with electronic health record and clinical decision support systems. Cloud computing facilitates AI development and deployment	Requires ongoing training for health care professionals to ensure effective use
Distributed surveillance	Analyzes data across geographically dispersed locations	Enables real-time population wide monitoring of disease trends. Helps track the spread of infectious pathogens and facilitates early intervention strategies	Inconsistent data formats, collection methods, and definitions across institutions can hinder accurate analysis; distributed systems are more vulnerable to cyberattacks
Federated learning	Trains models on data that is geographically or institutionally separated	Preserves data privacy by only sharing model updates, not the original data	Federated learning may lead to slightly less accurate models compared with traditional methods because of data heterogeneity; model inversion attacks pose security vulnerabilities; efficiently exchanging model updates creates communication overhead challenges

This table explores 3 key advancements in artificial intelligence (AI) that are transforming cardiology: Clinical AI operations (AI-Ops), distributed surveillance, and federated learning. Each advancement offers unique benefits for health care professionals and patients. Clinical AI-Ops streamlines the integration of AI tools into clinical workflows, whereas distributed surveillance enables large-scale disease monitoring across geographically dispersed locations. Federated learning tackles data privacy concerns by training AI models on separated data, promoting collaboration without compromising confidentiality. However, each approach also faces challenges, such as the need for ongoing training for clinical AI-Ops, data privacy, and quality issues in distributed surveillance, and potential accuracy limitations and communication overhead in federated learning. Overall, these advancements hold significant promise for the future of cardiology, paving the way for improved patient outcomes and population health management.

SHAP values provide transparent insights into individual predictions, elucidating the driving risk factors behind the model's decisions and facilitating clinical interpretation. It is, however, important to note that many of these strategies, when applied to algorithms types that are fundamentally not explainable only generate pseudoexplainability, and it remains to be determined whether from a regulatory point-of-view, true explainability and pseudoexplainability will be considered in the same way. Continued progress in XAI and seamless integration of this information (influential variables, bias measurement, calibration, etc) into clinical workflows will undoubtedly fortify the position of ML as an indispensable tool in the everyday practice of cardiology.

Table 4 comprehensively addresses analytical considerations in AI applications for cardiology, focusing on interpretability and explainability with specific AI methods such as LR, XGBoost, and a few other ML techniques discussed in this article. It highlights how these techniques contribute to feature importance, model performance metrics, complex model interpretation, and clinical validation, underscoring their role in enhancing transparency and reliability of AI-driven health care solutions.

Algorithmic Drift

Data drift, encompassing shifts in data distribution over time, and continuous integration and delivery play crucial roles in determining the deployment of ML applications.[58] As patient populations and health care practices evolve, the performance of deployed ML models may degrade. Monitoring algorithmic drift and retraining models with updated data becomes indispensable to sustain their effectiveness. Kore et al.[95] highlighted various causes of data drift, such as changes in user behaviour, shifts in the underlying data distribution, and changes in the environment in which the model operates. They investigated data drift within the context of chest radiograph disease classification. They used a dataset consisting of 239,235 chest radiographs (CXRs) and their associated imaging reports collected before and after the onset of the COVID-19 pandemic. The researchers employed a pretrained TorchXRayVision classifier, which had been fine-tuned to identify 14 different pathologies in CXR. They discussed 4 approaches to detect data drift, in which the first approach is to track model performance. The second and third approaches involve looking at the image data or model output, respectively. The fourth approach combines these 2 techniques. They also evaluated the approaches on synthetic data, in which they simulate changes in patient demographics and pathologies. Finally, they explore how the size of the data sample affects the sensitivity of the drift-detection approaches. Their results indicate that tracking performance of the model alone is not sufficient, and that data-drift detection is also pivotal. Furthermore, the study reveals a correlation between the sensitivity of drift detection and sample size, highlighting implications for timely intervention in clinical settings. Mitigation strategies include retraining the model,[96] monitoring the ML performance in combination with the risk factors distribution to monitor data drift.

Table 4. Analytical considerations, interpretability, and explainability in cardiology

Analytical consideration	Interpretability	Explainability	AI method
Feature importance	Methods such as SHAP and LIME provide insights into how each feature contributes to model predictions, enhancing transparency	Explainability techniques like SHAP values and LIME elucidate the influence of individual features on predictions, aiding in clinical decision making	SHAP, LIME, LR, XGBoost
Model performance metrics	Metrics such as AUC-ROC, precision, and recall are foundational for assessing model accuracy and reliability, contributing to the model's interpretability by quantifying its predictive power	Advanced metrics such as calibration plots and precision-recall curves go beyond traditional metrics to provide a nuanced understanding of model performance, enhancing the explainability of model outcomes	Calibration plots, precision-recall curves, LR, XGBoost
Complex model interpretation	Techniques such as UMAP and t-SNE enable visualization of high-dimensional data, facilitating intuitive interpretation of complex model outputs	UMAP and t-SNE help visualize complex data structures, aiding in the explanation of intricate model decisions and patterns to stakeholders	UMAP, t-SNE, XGBoost
Clinical validation and transparency	External validation and model transparency through detailed documentation and validation reports ensure that model decisions are understandable and replicable, supporting interpretability	Rigourous validation methodologies and transparent reporting mechanisms ensure that model outputs are explainable, fostering trust and adoption in clinical settings	External validation, transparent reporting mechanisms, LR, XGBoost

This table presents a comprehensive overview of how interpretability and explainability are addressed in various AI methods applied to cardiology. It outlines specific analytical considerations such as feature importance, model performance metrics, complex model interpretation, and clinical validation and transparency. For each consideration, the table details how different AI techniques such as LR, XGBoost, SHAP, LIME, UMAP, and t-SNE enhance both the understanding of the model's functioning (interpretability) and the insights into the model's decisions (explainability). This structured approach ensures that AI models used in cardiology are transparent, reliable, and effective in clinical decision making.

AI, artificial intelligence; AUC-ROC, area under the curve of the receiver operating characteristic; LR, logistic regression; LIME, local interpretable model-agnostic explanations; SHAP, SHapley Additive exPlanations; t-SNE, t-distributed stochastic neighbour embedding; UMAP, uniform manifold approximation and projection; XGB, extreme gradient boosting.

Distributed Surveillance and Federated Learning

Federated learning (FL), a specific type of distributed learning that allows training models on data that is geographically or institutionally separated.[97] This is particularly useful in health care when privacy concerns and regulations restrict data sharing. FL accomplishes this by training the model locally on each participant's data and only sharing the model updates, not the original data. These updates can be exchanged in a centralized or peer-to-peer manner. FL techniques are in constant evolution, including federated averaging and FL with differential privacy, to enhance model performance while preserving data privacy.[97-99]

Despite its potential, FL faces several challenges. Heterogeneity in data distributions among participating institutions can lead to model performance degradation.[99] Security vulnerabilities, such as model inversion attacks, threaten the privacy of sensitive data. In addition, communication overhead poses a significant challenge in exchanging model updates efficiently. Such techniques as data preprocessing and model aggregation help to alleviate heterogeneity issues. Differential privacy mechanisms protect against privacy breaches, whereas secure aggregation protocols mitigate security risks. Furthermore, optimizing communication protocols and federated optimization algorithms can reduce communication overhead.

Benefits of Digital Biomarkers for Personalized Medicine: Example of Predictive Allocation

Through careful selection and application of ML techniques tailored to the unique characteristics of cardiac data, researchers are paving the way for a transformative era in personalized cardiovascular medicine. Under the prevailing current clinical paradigm, rooted in EBM, personalized medicine prioritizes treatment based on the best available clinical evidence while acknowledging individual patient differences. Although some care can be stratified based on risk factors, true personalized care, guided by comprehensive patient characteristics and prediction models, is rare, despite its potential for better patient outcomes. A key unanswered question to health care is the potential benefit of personalized approaches for a broader patient population. To tackle this, we used data from multiple clinical trials to simulate the impact of personalized medicine.

Personalized Care Using Predictive Allocation

Guided by Predictive Approaches to Treatment effect Heterogeneity (PATH) guidelines,[100,101] we focused on predictive allocation, a method employing outcome models to assign treatment, aiming to minimize the probability of negative outcomes.[102] Computer simulations allow researchers to study complex systems, predict their behaviour under different conditions, and explore hypothetical scenarios. By applying prediction models to different trial groups within randomized clinical trials (RCTs), we aimed to estimate patient-specific outcomes under various treatment scenarios.[103] Comparing these simulations with standard treatment approaches will elucidate the potential population-level benefits of personalized medicine (Fig. 3). In our study,[104] we selected 3 RCTs from pediatric cardiology literature to

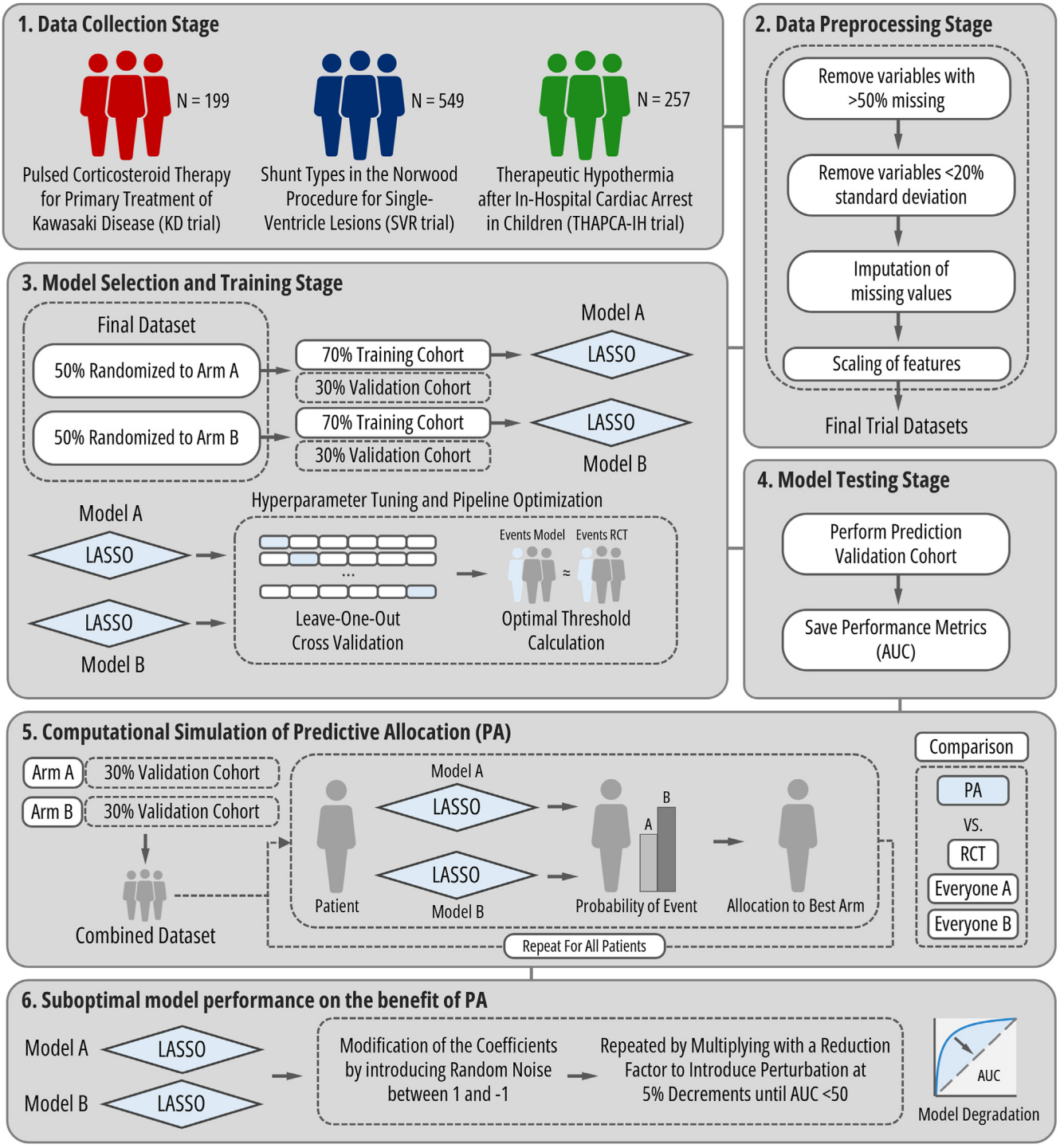

Figure 3. Personalized care using predictive allocation. The flowchart illustrates the process of personalized care through predictive allocation. First, prediction models are developed and optimized using training data from randomized clinical trials. Patients' likelihood of experiencing the primary outcome is predicted based on their characteristics and treatment arm. These predictions are used to compare treatment options for patients, aiming to assign each of them to the treatment with the lowest predicted probability of negative outcomes. The approach demonstrates significant reductions in predicted probabilities of primary outcomes, showcasing the potential benefits of personalized medicine. Reproduced from Jacquemyn X, Van den Eynde J, Chinni BK, Danford DM, Kutty S, Manlhiot C,[104] with permission from the Journal of the American Medical Informatics Association.

represent diverse clinical scenarios: the **P**ediatric **H**eart **N**etwork **S**ingle **V**entricle **R**econstruction (PHN SVR) trial,[105] the **T**herapeutic **H**ypothermia **A**fter **P**ediatric **C**ardiac **A**rrest-**I**n-**H**ospital (THAPCA-IH) trial,[106] and the Pediatric **H**eart **N**etwork **K**awasaki **D**isease (HNKD) trial. Within each RCT and treatment arm, patients were randomly divided into training and validation sets. LR models were then applied to each treatment-training group, and their performance was optimized. Using the predictive allocation approach, we used LR models to predict the likelihood of each patient experiencing the primary outcome, considering their characteristics and treatment arm. This allowed us to compare treatment options for each patient and assign them to the one with the lowest predicted probability. Our findings demonstrated that the use of predictive allocation could reduce the predicted probabilities of experiencing primary outcomes among patients in all 3 RCTs significantly, with reductions ranging from 15% to 35%.[104]

Conclusions

The integration of digital biomarkers from AI algorithms into cardiology is revolutionizing the field; enhancing diagnostic accuracy, prognosis, and risk assessment; and paving the way for precision and personalized medicine. Precision medicine focuses on targeting treatments based on patients' biological characteristics, In contrast, personalized medicine extends this approach by using ML and DL techniques to analyze complex and integrate high-dimensional clinical data from various sources, including ECG, echocardiograms, EHR, and wearable devices to create digital biomarkers. Personalized medicine also incorporates a broader array of data, including clinical, psychological, environmental, and socioeconomic dimensions. This approach allows for personalized diagnoses and tailored treatment plans optimized to maximize effectiveness and minimize adverse events. For example, recent advancements include the use of time-dependent models such as RNN for predicting progression of disease and early detection of conditions such as heart failure and arrhythmias. Also, the ability of AI to analyze cardiac imaging data, employing CNN for such tasks as segmentation in CT angiography and detection of valvular heart disease, has shown significant promise. Network analysis further enhances our understanding by mapping complex biological interactions, whereas multimodal learning frameworks integrate diverse data sources for comprehensive disease modelling.

Successful implementation of ML in cardiology requires meticulous methodology selection and consideration of various factors to ensure efficacy and reliability. Novel techniques such as feature permutation using Boruta and SHAP values provide insights into feature contributions, facilitating model interpretation and validation. Beyond accuracy, calibration techniques such as Platt scaling and isotonic regression refine risk predictions, bolstering the reliability of assessments. Addressing biases in ML models is crucial for generalizability. Strategies such as using diverse datasets, regularization, cross validation, ensemble learning, and multicentre data integration help mitigate these biases. XAI methods, including LIME and SHAP, enhance model transparency, fostering trust among clinicians and patients by elucidating the driving factors behind predictions.

Continuous monitoring of data drift, as highlighted in studies on CXR disease classification during the COVID-19 pandemic, underscores the need for ongoing model retraining and performance evaluation to sustain effectiveness. Novel implementation approaches, such as clinical AI-Ops, distributed surveillance, and FL, are transforming health care delivery. AI-Ops ensure seamless integration into clinical workflows, whereas distributed surveillance leverages AI for population-wide disease monitoring without compromising patient privacy. FL enables collaborative model training across institutions while preserving data confidentiality. In conclusion, AI is transforming cardiology by providing tailored health care solutions that account for the intricate and multifaceted nature of individual patient profiles. As AI technologies continue to evolve, they will play an increasingly integral role in shaping the future of health care, driving innovation and enhancing patient outcomes.

Ethics Statement

This is a review article and all research discussed in this article adhered to the applicable ethical guidelines.

Patient Consent

The authors confirm that patient consent in not applicable to this article. This article is a review of existing literature and as such does not qualify as human research, which would require patient consent.

Funding Sources

No funding was provided for this article.

Disclosures

The authors have no conflicts of interest to disclose.

References

1. Vo KT, Parsons DW, Seibel NL. Precision medicine in pediatric oncology. Surg Oncol Clin North Am 2020;29:63-72.

2. Duarte JD, Cavallari LH. Pharmacogenetics to guide cardiovascular drug therapy. Nat Rev Cardiol 2021;18:649-65.

3. Sackett DL, Rosenberg WMC. On the need for evidence-based medicine. J Public Health Med 1995;17:330-4.

4. Delpierre C, Lefèvre T. Precision and personalized medicine: what their current definition says and silences about the model of health they promote: implication for the development of personalized health. Front Sociol 2023;8:1112159.

5. Di Sanzo M, Cipolloni L, Borro M, et al. Clinical applications of personalized medicine: a new paradigm and challenge. Curr Pharm Biotechnol 2017;18:194-203.

6. Johnson KB, Wei WQ, Weeraratne D, et al. Precision medicine, AI, and the future of personalized health care. Clin Transl Sci 2021;14:86-93.

7. Hood L, Friend SH. Predictive, personalized, preventive, participatory (P4) cancer medicine. Nat Rev Clin Oncol 2011;8:184-7.

8. Howard J. Cognitive Errors and Diagnostic Mistakes: A Case-Based Guide to Criticial Thinking in Medicine. New York: Springer Publications, 2019:588.

9. Hearn J, Ross HJ, Mueller B, et al. Neural networks for prognostication of patients with heart failure. Circ Heart Fail 2018;11:e005193.

10. Bohr A, Memarzadeh K. The rise of artificial intelligence in healthcare applications. In: Bohr A, Memarzadeh K, eds. Artificial Intelligence in Healthcare. Cambridge: Academic Press, 2020:25-60.

11. Lysaght T, Lim HY, Xafis V, Ngiam KY. AI-assisted decision-making in healthcare: the application of an ethics framework for big data in health and research. Asian Bioeth Rev 2019;11:299-314.

12. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017;2:230-43.

13. Bozyel S, Simsek E, Kocyigit D, et al. Artificial intelligence-based clinical decision support systems in cardiovascular diseases. Anatol J Cardiol 2024;28:74-86.

14. Manlhiot C, van den Eynde J, Kutty S, Ross HJ. A primer on the present state and future prospects for machine learning and artificial intelligence applications in cardiology. Can J Cardiol 2022;38:169-84.

15. Zhang A, Xing L, Zou J, Wu JC. Shifting machine learning for healthcare from development to deployment and from models to data. Nat Biomed Eng 2022;6:1330-45.

16. Ben-Hur A, Weston J. A user's guide to support vector machines. Methods Mol Biol 2010;609:23-239.

17. Breiman L. Random forests. Machine Learning 2001;45:5-32.

18. Chen TQ, Guestrin C. XGBoost: a scalable tree boosting system, KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco: Association for Computing Machinery, 2016:785-94. Available at: https://doi.org/10.1145/2939672.2939785.

19. Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat 2001;29:1189-232.

20. Hancock JT, Khoshgoftaar TM. CatBoost for big data: an interdisciplinary review. J Big Data 2020;7.

21. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature 1986;323:533-6.

22. Bergstra J, Bengio Y. Random search for hyper-parameter optimization. J Machine Learning Res 2012;13:281-305.

23. Marcinkevics R, Vogt JE. Interpretable and explainable machine learning: a methods-centric overview with concrete examples. WIREs 2023;13:e1493.

24. Peng CYJ, Lee KL, Ingersoll GM. An introduction to logistic regression analysis and reporting. J Educ Res 2002;96:3-14.

25. James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning with Applications in R. New York: Springer Publications, 2013:1-14.

26. Hosmer DW, Lemeshow S, Sturdivant RX. Applied Logistic Regression. 3rd ed. Hoboken: Wiley, 2013:1-500.

27. Nygards ME, Hulting J. An automated system for ECG monitoring. Comput Biomed Res 1979;12:181-202.

28. Serhani MA, El Kassabi HT, Ismail H, Nujum Navaz A. ECG monitoring systems: review, architecture, processes, and key challenges. Sensors (Basel) 2020;20:1796.

29. Jahmunah V, Ng EYK, San TR, Acharya UR. Automated detection of coronary artery disease, myocardial infarction and congestive heart failure using GaborCNN model with ECG signals. Comput Biol Med 2021;134:104457.

30. Gandhi SO, Sabik L. Emergency department visit classification using the NYU algorithm. Am J Manag Care 2014;20:315-20.

31. Alowais SA, Alghamdi SS, Alsuhebany N, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ 2023;23:689.

32. Wells S, Furness S, Rafter N, et al. Integrated electronic decision support increases cardiovascular disease risk assessment four fold in routine primary care practice. Eur J Cardiovasc Prev Rehabil 2008;15:173-8.

33. Lin A, Manral N, McElhinney P, et al. Deep learning-enabled coronary CT angiography for plaque and stenosis quantification and cardiac risk prediction: an international multicentre study. Lancet Digit Health 2022;4:e256-65.

34. Abramoff MD, Whitestone N, Patnaik JL, et al. Autonomous artificial intelligence increases real-world specialist clinic productivity in a cluster-randomized trial. NPJ Digit Med 2023;6:184.

35. Farina JM, Pereyra M, Mahmoud AK, et al. Artificial intelligence-based prediction of cardiovascular diseases from chest radiography. J Imaging 2023;9:236.

36. Bourazana A, Xanthopoulos A, Briasoulis A, et al. Artificial intelligence in heart failure: friend or foe? Life (Basel) 2024;14:145.

37. Johnson KW, Torres Soto J, Glicksberg BS, et al. Artificial intelligence in cardiology. J Am Coll Cardiol 2018;71:2668-79.

38. Ledzinski L, Grzesk G. Artificial intelligence technologies in cardiology. J Cardiovasc Dev Dis 2023;10:202.

39. Feldner-Busztin D, Firbas Nisantzis P, Edmunds SJ, et al. Dealing with dimensionality: the application of machine learning to multi-omics data. Bioinformatics 2023;39:btad021.

40. Hewamalage H, Bergmeir C, Bandara K. Recurrent neural networks for time series forecasting: current status and future directions. Int J Forecasting 2021;37:388-427.

41. Chang YC, Wu SH, Tseng LM, Chao HL, Ko CH. AF detection by exploiting the spectral and temporal characteristics of ECG signals with the LSTM model. 2018 Computing in Cardiology Conference (CinC). Maastricht, Netherlands 2018;45:1-4.

42. Gupta U, Paluru N, Nankani D, Kulkarni K, Awasthi N. A comprehensive review on efficient artificial intelligence models for classification of abnormal cardiac rhythms using electrocardiograms. Heliyon 2024;10:e26787.

43. Zargarzadeh A, Javanshir E, Ghaffari A, Mosharkesh E, Anari B. Artificial intelligence in cardiovascular medicine: an updated review of the literature. J Cardiovasc Thorac Res 2023;15:204-9.

44. Kwon JM, Kim KH, Medina-Inojosa J, Jeon KH, Park J, Oh BH. Artificial intelligence for early prediction of pulmonary hypertension using electrocardiography. J Heart Lung Transplant 2020;39:805-14.

45. Yao X, Rushlow DR, Inselman JW, et al. Artificial intelligence-enabled electrocardiograms for identification of patients with low ejection fraction: a pragmatic, randomized clinical trial. Nat Med 2021;27:815-9.

46. Attia ZI, Kapa S, Lopez-Jimenez F, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. Nat Med 2019;25:70-4.

47. Attia ZI, Noseworthy PA, Lopez-Jimenez F, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients

with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. Lancet 2019;394:861-7.

48. Ko WY, Siontis KC, Attia ZI, et al. Detection of hypertrophic cardiomyopathy using a convolutional neural network-enabled electrocardiogram. J Am Coll Cardiol 2020;75:722-33.

49. Sakaue S, Hirata J, Kanai M, et al. Dimensionality reduction reveals fine-scale structure in the Japanese population with consequences for polygenic risk prediction. Nat Commun 2020;11:1569.

50. Yang Y, Sun H, Zhang Y, et al. Dimensionality reduction by UMAP reinforces sample heterogeneity analysis in bulk transcriptomic data. Cell Rep 2021;36:109442.

51. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. J Stat Software 2010;36:1-13.

52. Sarker IH. Machine learning: algorithms, real-world applications and research directions. SN Comput Sci 2021;2:160.

53. Eckhardt CM, Madjarova SJ, Williams RJ, et al. Unsupervised machine learning methods and emerging applications in healthcare. Knee Surg Sports Traumatol Arthrosc 2023;31:376-81.

54. Zhang Z, Murtagh F, Van Poucke S, Lin S, Lan P. Hierarchical cluster analysis in clinical research with heterogeneous study population: highlighting its visualization with R. Ann Transl Med 2017;5:75.

55. Murray EC, Delles C, Orzechowski P, et al. Vascular phenotypes in early hypertension. J Hum Hypertens 2023;37:898-906.

56. Verdonschot JAJ, Wang P, Derks KWJ, et al. Clustering of cardiac transcriptome profiles reveals unique: subgroups of dilated cardiomyopathy patients. JACC Basic Transl Sci 2023;8:406-18.

57. Nasios N, Bors AG. Variational learning for Gaussian mixture models. IEEE Trans Syst Man Cybern B Cybern 2006;36:849-62.

58. Sabry F, Eltaras T, Labda W, Alzoubi K, Malluhi Q. Machine learning for healthcare wearable devices: the big picture. J Healthc Eng 2022;2022:4653923.

59. Chen C, Qin C, Qiu H, et al. Deep learning for cardiac image segmentation: a review. Front Cardiovasc Med 2020;7:25.

60. Bui V, Hsu LY, Chang LC, et al. DeepHeartCT: a fully automatic artificial intelligence hybrid framework based on convolutional neural network and multi-atlas segmentation for multi-structure cardiac computed tomography angiography image segmentation. Front Artif Intell 2022;5:1059007.

61. Alkhodari M, Fraiwan L. Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings. Comput Methods Programs Biomed 2021;200:105940.

62. Sonawane AR, Weiss ST, Glass K, Sharma A. Network medicine in the age of biomedical big data. Front Genet 2019;10:294.

63. Sonawane AR, Aikawa E, Aikawa M. Connections for matters of the heart: network medicine in cardiovascular diseases. Front Cardiovasc Med 2022;9:873582.

64. Valenzuela JF, Monterola C, Tong VJC, Ng TP, Larbi A. Health and disease phenotyping in old age using a cluster network analysis. Sci Rep 2017;7:15608.

65. Pepe G, Appierdo R, Ausiello G, Helmer-Citterich M, Gherardini PF. A meta-analysis approach to gene regulatory network inference identifies key regulators of cardiovascular diseases. Int J Mol Sci 2024;25:4224.

66. Acosta JN, Falcone GJ, Rajpurkar P, Topol EJ. Multimodal biomedical AI. Nat Med 2022;28:1773-84.

67. Kline A, Wang H, Li Y, et al. Multimodal machine learning in precision health: a scoping review. NPJ Digit Med 2022;5:171.

68. Bullock-Palmer RP, Rosario KF, Douglas PS, et al. Multimodality cardiac imaging and the imaging workforce in the United States: diversity, disparities, and future directions. Circ Cardiovasc Imaging 2024;17:e016409.

69. Milosevic M, Jin Q, Singh A, Amal S. Applications of AI in multimodal imaging for cardiovascular disease. Front Radiol 2023;3:1294068.

70. Rim B, Sung NJ, Min S, Hong M. Deep learning in physiological signal data: a survey. Sensors (Basel) 2020;20:969.

71. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. J Am Med Inform Assoc 2017;24:361-70.

72. Minic A, Jovanovic L, Bacanin N, et al. Applying recurrent neural networks for anomaly detection in electrocardiogram sensor data. Sensors (Basel) 2023;23:9878.

73. Guk K, Han G, Lim J, et al. Evolution of wearable devices with real-time disease monitoring for personalized healthcare. Nanomaterials (Basel) 2019;9:813.

74. Bayoumy K, Gaber M, Elshafeey A, et al. Smart wearable devices in cardiovascular care: where we are and how to move forward. Nat Rev Cardiol 2021;18:581-99.

75. Tan L, Yu K, Bashir AK, et al. Toward real-time and efficient cardiovascular monitoring for COVID-19 patients by 5G-enabled wearable medical devices: a deep learning approach. Neural Comput Appl 2023;35:13921-34.

76. Hughes A, Shandhi MMH, Master H, Dunn J, Brittain E. Wearable devices in cardiovascular medicine. Circ Res 2023;132:652-70.

77. Steyerberg EW. Introduction. In: Steyerberg EW, ed. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Berlin: Springer International Publishing, 2019.

78. Pessach D, Shmueli E. Algorithmic fairness. In: Rokach L, Maimon O, Shmueli E, eds. Machine Learning for Data Science Handbook: Data Mining and Knowledge Discovery Handbook. Berlin: Springer International Publishing, 2023:867-86.

79. Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. A survey on bias and fairness in machine learning. ACM Comput Surv 2021;54:15.

80. Xu J, Xiao Y, Wang WH, et al. Algorithmic fairness in computational medicine. EBioMedicine 2022;84:104250.

81. Giovanola B, Tiribelli S. Beyond bias and discrimination: redefining the AI ethics principle of fairness in healthcare machine-learning algorithms. AI Soc 2023;38:549-63.

82. McCradden MD, Joshi S, Mazwi M, Anderson JA. Ethical limitations of algorithmic fairness solutions in health care machine learning. Lancet Digit Health 2020;2:E221-3.

83. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. Ann Intern Med 2018;169:866-72.

84. Chen RJ, Wang JJ, Williamson DFK, et al. Algorithmic fairness in artificial intelligence for medicine and healthcare. Nat Biomed Eng 2023;7:719-42.

85. Chen E, Prakash S, Reddi VJ, Kim D, Rajpurkar P. A framework for integrating artificial intelligence for clinical care with continuous

therapeutic monitoring [e-pub ahead of print]. Nat Biomed Eng 2023. https://doi.org/10.1038/s41551-023-01115-0.

86. Bhagat SV, Kanyal D. Navigating the future: the transformative impact of artificial intelligence on hospital management: a comprehensive review. Cureus J Med Scie 2024;16:e54518.

87. Koppad S, Annappa B, Gkoutos GV, Acharjee A. Cloud computing enabled big multi-omics data analytics. Bioinform Biology Insights 2021;15:11779322211035921.

88. Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell 2020;2:56-67.

89. Huang YX, Li WT, Macheret F, Gabriel RA, Ohno-Machado L. A tutorial on calibration measurements and calibration models for clinical prediction models. J Am Med Inform Assoc 2020;27:621-33.

90. Jiang XQ, Osl M, Kim J, Ohno-Machado L. Calibrating predictive model estimates to support personalized medicine. J Am Med Inform Assoc 2012;19:263-74.

91. Nazer LH, Zatarah R, Waldrip S, et al. Bias in artificial intelligence algorithms and recommendations for mitigation. PLOS Digit Health 2023;2:e0000278.

92. Amiri SS, Mottahedi S, Lee ER, Hoque S. Peeking inside the black-box: explainable machine learning applied to household transportation energy consumption. Comput Environ Urban Syst 2021;88:101647.

93. de Sousa IP, Vellasco MMBR, da Silva EC. Local interpretable model: agnostic explanations for classification of lymph node metastases. Sensors (Basel) 2019;19:2969.

94. Ribeiro MT, Singh S, Guestrin C. Why Should I Trust You?, Explaining the Predictions of Any Classifier. KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco: Association for Computing Machinery, 2016:1135-44.

95. Kore A, Abbasi Bavil E, Subasri V, et al. Empirical data drift detection experiments on real-world medical imaging data. Nat Commun 2024;15:1887.

96. Rahmani K, Thapa R, Tsou PL, et al. Assessing the effects of data drift on the performance of machine learning models used in clinical sepsis prediction. Int J Med Inform 2023;173:104930.

97. Rieke N, Hancox J, Li W, et al. The future of digital health with federated learning. NPJ Digit Med 2020;3:119.

98. Xu J, Glicksberg BS, Su C, Walker P, Bian J, Wang F. Federated learning for healthcare informatics. J Healthc Inform Res 2021;5:1-19.

99. Choi G, Cha WC, Lee SU, Shin SY. Survey of medical applications of federated learning. Healthc Inform Res 2024;30:3-15.

100. Kent DM, Paulus JK, van Klaveren D, et al. The predictive approaches to treatment effect heterogeneity (PATH) statement. Ann Intern Med 2020;172:35-45.

101. Rekkas A, Paulus JK, Raman G, et al. Predictive approaches to heterogeneous treatment effects: a scoping review. BMC Med Res Methodol 2020;20:264.

102. Zhang B, Tsiatis AA, Davidian M, Zhang M, Laber E. Estimating optimal treatment regimes from a classification perspective. Statistics 2012;1:103-14.

103. Rekkas A, Rijnbeek PR, Kent DM, Steyerberg EW, van Klaveren D. Estimating individualized treatment effects from randomized controlled trials: a simulation study to compare risk-based approaches. BMC Med Res Methodol 2023;23:74.

104. Jacquemyn X, Van den Eynde J, Chinni BK, Danford DM, Kutty S, Manlhiot C. Computational simulation of the potential improvement in clinical outcomes of cardiovascular diseases with the use of a personalized predictive medicine approach. J Am Med Inform Assoc 2024;31:1704-13.

105. Ohye RG, Sleeper LA, Mahony L, et al; Pediatric Heart Network I. Comparison of shunt types in the Norwood procedure for single-ventricle lesions. N Engl J Med 2010;362:1980-92.

106. Moler FW, Silverstein FS, Holubkov R, et al; THAPCA Investigators. Therapeutic hypothermia after in-hospital cardiac arrest in children. N Engl J Med 2017;376:318-29.