# Artificial Intelligence for Risk Assessment on Primary Prevention of Coronary Artery Disease

Shang-Fu Chen[1,2] · Salvatore Loguercio[1,2] · Kai-Yu Chen[1,2] · Sang Eun Lee[1,2,3] · Jun-Bean Park[1,2,4,5] · Shuchen Liu[1,2] · Hossein Javedani Sadaei[1,2] · Ali Torkamani[1,2]

## Abstract

**Purpose of Review** Coronary artery disease (CAD) is a common and etiologically complex disease worldwide. Current guidelines for primary prevention, or the prevention of a first acute event, include relatively simple risk assessment and leave substantial room for improvement both for risk ascertainment and selection of prevention strategies. Here, we review how advances in big data and predictive modeling foreshadow a promising future of improved risk assessment and precision medicine for CAD.

**Recent Findings** Artificial intelligence (AI) has improved the utility of high dimensional data, providing an opportunity to better understand the interplay between numerous CAD risk factors. Beyond applications of AI in cardiac imaging, the vanguard application of AI in healthcare, recent translational research is also revealing a promising path for AI in multi-modal risk prediction using standard biomarkers, genetic and other omics technologies, a variety of biosensors, and unstructured data from electronic health records (EHRs). However, gaps remain in clinical validation of AI models, most notably in the actionability of complex risk prediction for more precise therapeutic interventions.

**Summary** The recent availability of nation-scale biobank datasets has provided a tremendous opportunity to richly characterize longitudinal health trajectories using health data collected at home, at laboratories, and through clinic visits. The ever-growing availability of deep genotype-phenotype data is poised to drive a transition from simple risk prediction algorithms to complex, "data-hungry," AI models in clinical decision-making. While AI models provide the means to incorporate essentially all risk factors into comprehensive risk prediction frameworks, there remains a need to wrap these predictions in interpretable frameworks that map to our understanding of underlying biological mechanisms and associated personalized intervention. This review explores recent advances in the role of machine learning and AI in CAD primary prevention and highlights current strengths as well as limitations mediating potential future applications.

**Keywords** Artificial intelligence · Coronary artery disease · Primary prevention

✉ Ali Torkamani
  atorkama@scripps.edu

1 Scripps Research Translational Institute, La Jolla, CA 92037, USA

2 Department of Integrative Structural and Computational Biology, Scripps Research, La Jolla, CA 92037, USA

3 Department of Cardiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul 05505, South Korea

4 Division of Cardiology, Department of Internal Medicine, Seoul National University College of Medicine, Seoul 03080, South Korea

5 Cardiovascular Center, Seoul National University Hospital, Seoul 03080, South Korea

## Coronary Artery Disease

CAD is the leading cause of mortality and morbidity in the USA and Europe and among the most prevalent and severe manifestation of cardiovascular disease [1, 2]. CAD is characterized by atherosclerotic lesions, whereby plaques consisting of fatty deposits, inflammatory white blood cells, and smooth muscle cells accumulate and obstruct the coronary arteries [3]. Early symptoms of obstructive CAD due to vascular stenosis can result in angina, arrhythmias, and transient ischemic attacks [4]. Progressive intracoronary thrombosis with plaque rupture can eventually develop into a complete blockage, leading to myocardial infarction, heart failure, and sudden cardiac death [5]. Other vascular diseases are also

common in patients with CAD, such as carotid atherosclerosis, peripheral arterial disease, and stroke [4]. CAD is a complex multifactorial disease with nearly 300 risk factors statistically associated with its development [6–8]. CAD also shows significant heterogeneity across geographic regions, which makes generalized early diagnosis difficult to achieve. Despite WHO Member States global action plan for the prevention and control of CAD, the prevalence of CAD and CAD-related healthcare costs have continued to increase [9–11]. Thus, there continues to be a pressing need to build an early prevention eco-system to reduce the global public health burden of CAD.

## CAD Risk Factors

CAD risk factors can be divided into (1) modifiable risk factors (including hypertension, hyperlipidemia, hyperuricemia, diabetes mellitus, obesity, smoking, psychosocial stress, diet, sedentary lifestyle, and socioeconomic status), (2) non-modifiable risk factors (including age, gender, and genetic factors), and (3) risk-enhancing comorbid disease factors (including non-alcoholic fatty liver disease (NAFLD), chronic kidney disease (CKD), systemic lupus erythematosus (SLE), rheumatoid arthritis (RA), inflammatory bowel disease (IBD), hypercoagulability, human immunodeficiency (HIV), thyroid disease). Some of these risk factors can be captured by laboratory measurements, indices or biomarkers for daily clinical practice, and target for treatments — including vital signs (body temperature, pulse rate, respiration rate, heartbeat rate, blood pressure), body mass index (BMI), ankle-brachial index (ABI), coagulation indexes, cardiac calcium score, lipids/lipoproteins (LDL cholesterol, HDL cholesterol, total cholesterol triglycerides, lipoprotein(a), homocysteine), high-sensitivity c-reactive protein (hsCRP), and inflammatory cytokines.

## Primary Prevention of CAD

Current primary prevention guidelines in the USA and UK involve the use of additive risk assessment tools (including Reynolds score, Framingham risk score, pooled cohort risk equations (PCE), and QRISK), each including some of the risk factors above, and generally assigning individuals into low-, intermediate-, and high-risk populations. General recommendations for primary prevention include "Life's Simple 8": to get active, acquire adequate sleep, eat better, lose weight, stop smoking, control cholesterol, manage blood pressure, and reduce blood sugar. For at-risk populations identified by guideline algorithms, the first-line intervention remains lifestyle modification including smoking cessation, Mediterranean diet, intentional weight loss, and increasing physical activity. Guidelines typically recommend medications for individuals considered to be at higher

risk or carrying the presence of abnormal levels of specific biomarkers. The mainstay therapy in primary prevention is lipid lowering with statins, ezetimibe, and PCSK9 inhibitors. Other major biomarkers targeted for primary prevention of CAD are controlled via hypertension management (blood pressure lowering agents) and diabetes management (SGLT2 inhibitors, GLP1 receptor agonists). Numerous other contributory conditions and corresponding primary prevention approaches included lipidome remodeling (N-3 fatty acids), gut microbiome remodeling (small molecules, prebiotics, probiotics, or cyclic peptides [12]), antiplatelet (aspirin), anticoagulation (low-dose rivaroxaban), anti-inflammatory (colchicine), and vaccination (influenza and COVID-19) or PrEP (HIV). Multiple biomarkers may be targeted with a single therapeutic agent with polypills.

## Emerging Opportunities for AI in Early CAD Risk Assessment

AI models provide an opportunity to combine CAD risk factors into more complex risk assessment models, empowering physicians to make clinical decisions by harnessing the wealth of available health information for each individual [13]. There is not much dispute about the potential value of predictive models in cardiology, especially in early CVD detection [13]. Numerous studies have indicated that the performance of machine learning (ML)-based risk assessment models may exceed traditional risk assessments, even when simply using well-established cardiovascular disease (CVD) risk factors [14–16]. Moreover, data with different modalities, e.g., ECGs, chest X-rays, laboratory values, and polygenic risk scores (PRS), can be harnessed in these models to drive multi-modal precision CAD prevention [17]. When combined with genetic data, risk assessments can be made earlier, potentially leading to improved primary prevention [18]. Through genetic insights, we can measure the impact of familial ties on CVD and detect those predisposed to risk well ahead of the primary indicators of atherogenesis [19].

Most learning applications are achieved through supervised learning approaches, requiring labeled ground truth training data [20]. However, many of the available supervised learning algorithms (including generalized linear models (GLM), support vector machines (SVM), or decision trees) contend with the bias-variance tradeoff [21]. This tradeoff represents a situation where a model may overly adapt to its training data, known as overfitting, or conversely, may be too generalized, thereby missing intricate data patterns and resulting in underfitting [22]. Ensemble prediction models, by their very nature, attempt to navigate this tradeoff by amalgamating various algorithms or utilizing multiple iterations of a single algorithm through several forest and boosting methods (AdaBoost,

LightGBM, CatBoost, and XGBoost), aiming to strike an optimal balance between bias and variance [23]. Further, modern computing hardware has revitalized a subfield of AI inspired from biological neuron connectivity — deep learning (DL) — which now includes novel neural-network architectures successful in different domains, namely, convolutional neural networks (CNNs) for image recognition, recurrent neural networks (RNNs) for time series forecasting, and attention-based models for natural language processing (NLP) including large language models (LLM) [24, 25]. Finally, unsupervised AI methods, including some DL methods, can also enable improved clinical diagnosis of CAD by learning representative patterns free from human hypotheses in order to capture cryptic early symptoms from high dimensional data [26, 27].

At a high level, these major predictive modeling approaches can be applied to combine CAD risk factors into AI models in the following ways: (1) combining traditional biomarkers in AI models, (2) integrating additional genetic and other omic risk factors into more comprehensive risk assessment models, (3) including sensor-based feeds for real-time risk detection, (4) integrating various imaging modalities for active disease detection, and (5) capturing other data from EHRs using AI. In the remainder of this review, we will discuss the recent specific applications of AI for CAD primary prevention, considering each class of CAD risk factor, and provide our view on the necessary future iterations of these approaches in order to produce actionable insights linking causal mechanisms to preventive interventions (Fig. 1).

## Laboratory Biomarker Risk Assessment with AI

Simply combining traditional risk factors or contemporary risk tools into more complex predictive frameworks has provided evidence of low-hanging fruit for AI models in CAD risk assessment [16, 28–33]. For example, Petrazzini et al. built an EHR score from tabular clinical features with ML framework (random forest, gradient boosted trees, SVM stacked model) and improved CAD prediction from ASCVD score by 12% in the BioMe Biobank and 9% in the UK Biobank [34]. Further gains in CAD risk assessment accuracy have also been described with ensemble prediction models [35]. ML approaches have also facilitated novel biomarker identification and prioritization, including purine-related metabolites [36], apolipoprotein B [37], glutathione peroxidase-3 [38], epicardial adipose tissue [39], sleep heart rate variability [40], plasma lipids (184 lipids in lipidome) [41], and serum sphingolipids [42]. AI models can also be used to impute biomarker levels when direct determination is not available [43–46]. The success of these stack and/or ensemble prediction models suggest multiple health trajectories leading to CAD, with actionable information potentially more accurately captured by these complex models relative to the more simple, traditional, linear risk scoring approaches [47].
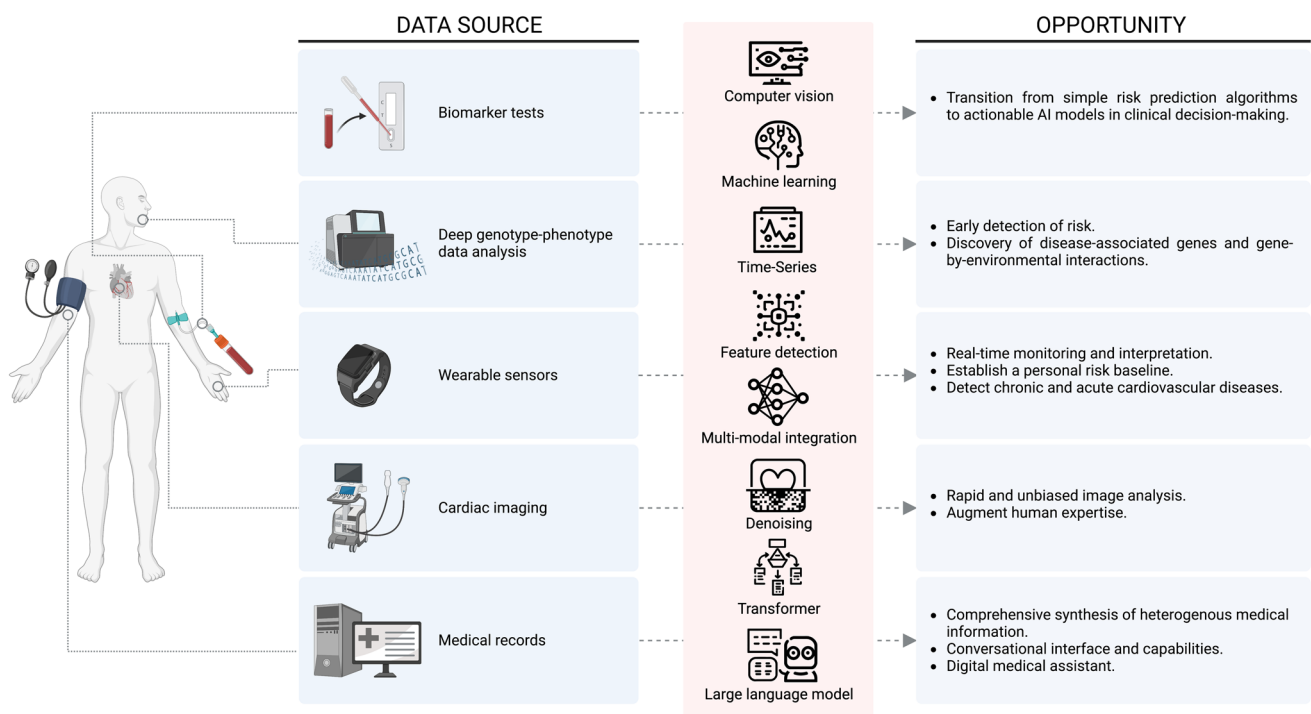


**Fig. 1** Opportunities for AI-driven CAD prevention and management

## Genetically Informed Risk Assessment Models

CAD risk has a strong genetic component driven by the interplay of environmental factors with genetic susceptibility factors ranging from monogenic (Mendelian) to highly polygenic risk [19]. Twin studies suggest a heritability of 50–60% for fatal CAD [48, 49]. Genetic risk assessment based on germline DNA provides a robust orthogonal predictor to laboratory biomarker-based risk factors and allows for early risk screening before other clinical measurements become informative [50–56]. CAD polygenic risk scores ranging from dozens to thousands of common variants can convey risk explaining between ~ 10% [57, 58] and ~ 40% [7, 59, 60] of disease heritability. A high CAD PRS is associated with increased benefit from lipid lowering interventions, including both statins [61–63] and PCSK9 inhibitors [64, 65]. These observations provide an opportunity to prioritize lipid lowering interventions for individuals predicted to receive improved benefit in the context of a standard guideline-based risk assessment framework [60, 66–69]. Linear and non-linear combinations of multiple PRSs have also been demonstrated to improve the polygenic prediction [70–72]. However, it is currently unclear how to best combine genetic information with clinical risk factors to demonstrate significant clinical benefit in large cohorts [55, 73–76].

AI approaches have been applied to GWAS data for a number of purposes including the identification of novel prognostic/causal markers and druggable targets [77–80]. ML/DL models for systematic post-GWAS analysis include those for functional annotation, functional fine-mapping, or functional scoring (pathogenicity or cell-specific importance scoring) to infer underlying regulatory mechanisms of non-coding CAD risk loci [81–87], and improving of the accuracy and cross-ancestry transferability of CAD-PRS [88, 89]. Other examples incorporate biological networks to further insights, like GCN-GENE, a DL model which leveraged propagation of GWAS signals in biological networks to identify additional CAD-related genes [90], or GenNet [91], another approach leveraging biological networks to perform genotype-to-phenotype mapping and improving risk assessment.

These risk factors, separated into biological pathways, interact with the environment in differing ways and may be further amenable to tree-based learning methods for the construction of genetic risk models theoretically able to capture differing gene-by-environmental interactions across diverse populations [79, 92–94]. Despite evidence of complex interactions between genetics and environment in mediating CAD risk [95, 96], existing linear risk prediction models including genetic risk often do not capture gene–gene or gene-environmental interactions in risk assessment [97–100]. AI models provide an opportunity to capture these complex relationships. Forrest et al. implemented a random forest-based ML system to this end for the cross-sectional detection of CAD

and achieved AUROC as 0.89 [101]. Nam et al. used a semi-supervised multi-layered network to devise a network-based MI risk score including interactions between PheWAS and PRS features, achieving an AUC improvement of 28.29% compared to PRS-alone model [102]. Steinfeldt et al. presented NeuralCVD from a deep survival machine to outperform linear algorithms (SCORE, ASCVD, QRISK3, and linear cox model) and identified an interaction of clinical risks and PRSs [103]. However, few studies have succeeded in applying these concepts in large-scale, longitudinal, incident CAD risk prediction [104]. To optimize the efficacy of these models in primary prevention, rigorous prospective studies with guideline recommendations are vital, as they hold the potential to both elucidate incident risk predictions and guide evidence-based preventive interventions [105–107].

## Multi-omics Data for Precision Risk Assessment

High-throughput transcriptomic, epigenetic, proteomic, and genomic technologies have enabled other comprehensive biomarker surveys in CAD [108–110]. These analytes provide signatures of the current physiological state, which may then be used to diagnose CAD or provide predictions of risk for future events. For example, transcriptomic profiles can be used to detect regulatory signatures that capture the current biological or pathological state of tissue. Few studies have progressed beyond cross-sectional detection of potential biomarkers to the prospective prediction of CAD events and myonecrosis. One exception includes transcriptomic predictors of impeding acute myocardial infarction (AMI) events from whole blood-derived circulating endothelial cells (CECs) [111–114]. The CEC transcriptomic signature potentially represents the biological state at the site of plaque rupture, serving as a biomarker of its present state, and potentially predictive of risk earlier than conventional biomarkers such as troponin and CK-MB. Another example of the utility of transcriptomic data includes a recent study which constructed a single-nucleus atlas of chromatin accessibility in human coronary arteries and identified specific cell-type regulatory mechanisms. By employing a statistical genetics and ML strategy, the study prioritized candidate regulatory variants and mechanisms for CAD loci and revealed detailed mechanisms connecting cell types, causal genes, and CAD risk variants in diverse populations [87].

Additionally, since CAD risk varies in relation to diverse exogenous and endogenous factors such as environment, diet, and lifestyle, DNA methylation signatures associated with CAD may be a proxy for exogenous exposures over time. Thus far, DNA methylation, integrated with genetic and clinical features, has been used to predict the risk of early heart failure, large-artery atherosclerosis stroke, and the risk of CAD via supervised learning with some success. However, few of these studies have demonstrated

generalizability in diverse and independent datasets. One study developed an ensemble predictor for incident CAD risk in different cohorts and combined them using a cross-study learning approach. This study demonstrated the feasibility of a genome-wide, epigenomic risk score for the prediction of future CAD events, possibly in individuals who would not be identified by other conventional risk metrics, but true clinical utility of epigenome risk scores remains to be demonstrated [100, 115, 116].

Proteomic studies appear to have similar value in situations where traditional risk assessment methods have limited predictive power for certain high-risk individuals. These scenarios include those with high risk of recurrent atherosclerotic events that require more intensive therapeutic interventions, individuals with known CAD risk but with well-controlled LDL-cholesterol and blood pressure, as well as those with multiple chronic conditions. In one example, researchers developed prognostic risk models based on plasma proteomics coupled with AI that can better predict cardiovascular outcomes within a relatively short period of time [117]. It was also suggested that the plasma proteome predictor could act as a universal surrogate endpoint for CAD, providing an avenue to improve patient outcomes through selective drug allocation and better monitoring in phase 2 clinical trials. In addition, plasma proteomics has been used to develop models for survival prediction after AMI [118, 119], prediction of recurrent events [120, 121], and improved assessment of risk for primary events [122].

Finally, dysregulation of gut microbiome has been shown to be associated with many chronic inflammatory diseases and is connected to the emergence and progression of several CAD-related risk factors [123]. However, it remains a major challenge to disentangle the possible impact of metabolic disruptions on a dysfunctional gut microbiota vs an imbalanced gut microbiota that more causally drives pathogenesis, potentially exacerbating ischemic heart disease processes at a later stages [124–126]. AI strategies such as shallow learning algorithms (random forest, support vector machine, and neural network etc.) are generally only useful for classification and identification of taxonomic differences between healthy and diseased individuals [123, 127]. They are not yet able to disentangle the interplay of reactive vs potentially causal changes that are observed in an abnormal microbiome.

## Real-Time Sensor-Based Risk Monitoring for Early Detection and Prevention of CAD Using AI

While biomarkers provide intermittent snapshots of health, smart medical devices and biosensors have the potential to revolutionize CAD risk monitoring by allowing real-time, longitudinal, collection of risk factor information

and trajectory estimation [128, 129]. Wireless networks, remote data centers, and edge computing enable wearables to monitor risk factors in real time [130, 131]. This eventual internet of things (IoT) also includes AI-assisted wearables that promise to provide accurate point-of-care diagnosis [132, 133], and will eventually cross-over into omics and laboratory-based biomarkers through the use of biochemical sensors. Initially, by building upon baseline applications of portable sensors for the automatic monitoring of cardiac rhythm disturbances, future iterations of cardiac biosensors could detect acute cardiovascular events and longitudinal factors to build up a personal risk baseline [128, 134–138]. In this section, we review recent applications and provide future perspectives on the use of various digital health and biosensor devices for CAD prevention with state-of-art AI approaches [139].

## Heart Rhythm (Cardiac Signal) Monitoring

Central to CAD is the heart and the change of mechanical, colorimetric, and electrical signals it produces with each heartbeat. These physical signals can be measured by electrodes, optical sensors, or motion sensors and interpreted into various biosignals including electrocardiogram (ECG), photoplethysmogram (PPG), stereocardiogram (SCG), phonocardiogram (PCG), ballistocardiogram (BCG), gyrocardiogram (GCG), or impulse cardiogram (ICG) for cardiac monitoring [140]. These signals can in turn be used to detect cardiac risks like atrial fibrillation and/or other heart rhythm disturbances indicative of future disease risk or pathology [141]. For example, He et al. extracted features of 30 dimensions from PPG signals to assess hemorrhagic risk in patients with CAD using an XGBoost regression model and achieving an AUC of 0.76 with tenfold cross validation [142]. Neural network-based models promise to improve sensor based prediction through their ability to denoise, annotate, and perform feature extraction on time traces [143–147]. ECG, PPG, and SCG can be used for remote heart condition monitoring when deployed using wearable devices, especially during exercise [148]. And with AI-assisted early abnormal signal detection, cardiac signals can be further used for CAD diagnosis and prediction [149–152]. While many of these algorithms operate on 12-lead ECG data, some researchers have demonstrated comparable performance with single-lead ECGs [152–154]. Despite advances in detecting specific cardiovascular abnormalities, the full utility of sensor-based signals for long-term risk prediction remains unproven. Addressing challenges such as confounding variables, temporal ambiguity, distinguishing genuine signals from noise, and population variability would be a further obstacle to overcome.

## Physical Activity Monitoring

Although lifestyle can be difficult to ascertain in a single clinic visit, it is been identified as mediating several risk factors for CAD. Lifestyle risk factors can be measured via accelerometer-based step counting, energy expenditure [155], sleeping composites [156, 157], and other factors [120, 139]. These activity metrics can then be included in ML models for prediction of cardiac outcomes. For instance, Nguyen et al. measured accelerometer-measured daily total movements and found an association between this measure of physical activity and risk of incident CVD in women and young adult [158, 159]. Triaxial accelerometer-based physical activity data has been used to demonstrate CAD risk factor reduction by changing chronoactivity [160, 161]. And Huang et al. designed an ensemble ML algorithm for the prediction of coronary artery calcium from predictors including several lifestyle and physical activity features [162]. Despite the breakthrough in AI and wearable technology [163, 164], the current studies in the field are limited to benefit with cardiac telerehabilitation in patients with CAD [165, 166]. Actual, real-time, and useful inclusion of physical activity measures in prediction of incident CAD risk has yet to be demonstrated.

## Biochemical Sensors

Biochemical sensors transform a biochemical analyte into an electronic signal, often using an integrated optical, acoustic, magnetic, or electrochemical sensing array operating in biofluids like blood, sweat, saliva, or urine. This technology can be used to conduct non-invasive, cost-effective, multi-analytes scans of human metabolites, and in quick response to lifestyle influences [167]. These scans have been used in CAD prediction models, integrating multi-modal signals of BP, temperature, ECG, glucose, hemoglobin, and oxygen levels, and achieving 97% accuracy with a minimum redundancy maximum relevance feature selection [168]. Multiple vital signs including electrodermal activity can also be combined in ML models to aid in the detection of sleep stages, which can then be used to improve CAD risk prediction [169, 170]. More unusual examples include the use of chemical gas sensors for the detection of CAD risk via an electronic nose [171]. Or novel quantum sensing approaches used to detect cardiac amyloidosis [172]. Biochemical sensing is an up-and-coming application area where clinical validation is still pending [173]. Economical deployment of these sensors and a system of interpretation and alerts will be the major challenge to overcome with these information-rich technologies.

## Environmental Sensing

Sensing can extend beyond the body and range from local neighborhood environmental monitoring [174], to include larger structural elements of society [175]. Several cohort studies have demonstrated the value of wearable sensors by capturing the complex gene-environment interactions [176] as well as the impact of longitudinal actionable changes [177]. One promising area of research involves the use of GPS technology and other environmental sensors to collect real-time data on an individual's exposure to air pollution, water quality, and other environmental factors that may increase their risk of CAD [178]. This can be valuable in developing personalized prevention strategies that consider an individual's unique travel patterns and contextual background [179, 180]. In addition to environmental factors, social factors play an important role in the development of CAD, such as social interaction and community engagement [181]. Integrating this data potentially enables AI-powered wearables to recognize regions with heightened environmental risk factors and provide tailored lifestyle coaching to reduce the incidence of CAD through sustainable behavioral changes [182, 183]. These applications are likely to be further into the future as a deeper understanding of the interplay between endogenous and environmental risk factors must be appreciated before these datastreams extend to individual level vs population level utility.

# Advanced Applications of AI in Noninvasive Imaging for CAD Risk Evaluation

AI has been increasingly applied to cardiovascular imaging for risk stratification of CAD, by virtue of its ability to accurately quantify prognostic biomarkers from image data, in addition to the reduction in cost and improvement in image acquisition and interpretation. This section summarizes recent promising applications of AI across various noninvasive imaging modalities, including coronary artery calcium imaging, coronary computed tomography (CT) angiography, peri-coronary/epicardial adipose tissue imaging, nuclear imaging, and retinal imaging, for the improved risk assessment of CAD, which can better guide decision-making in the primary prevention of CAD.

## Coronary Artery Calcium Scoring

As coronary artery calcium (CAC) is a highly specific feature of atherosclerosis, CAC scoring (CACS) has emerged as a powerful and widely available means of predicting risk for atherosclerotic cardiovascular diseases, particularly useful for guiding primary prevention therapy decisions [184–186]. AI approaches have gained great attention due

to their promising automation capabilities for annotation of the calcified lesions. Sandstedt et al. used a ML model that integrates patient-specific heart-centric coordinates, local voxel images, and the coronary territory map to evaluate the diagnostic efficacy of this AI-driven, automated CACS software against semi-automated software, using the same ECG-gated CT images [187]. In this study, the AI-based method was less time-demanding, with an excellent correlation and agreement with the semi-automated one. The limited accessibility of ECG-gated CT represents a critical issue restricting its routine use. The millions of people undergo routine chest CT scans and demonstrate CAC, but its quantitation is not feasible yet. In this regard, recent studies have focused on the application of convolutional neural networks (CNN) to a wide range of CT examination types, including low-dose chest CT and radiation therapy planning CT, and suggest that AI-based CACS quantification is robust across the different CT protocols [188, 189]. The future application of AI-based CAC assessment is the microcalcification quantification, since the current methods are limited to advanced calcification. Considering that microcalcifications can induce to vascular stiffening and plaque rupture [190], AI-based microcalcification quantification can allow further risk stratification in primary prevention of individuals with normal results by conventional imaging.

## Coronary CT Angiography

Coronary CT angiography (CCTA) is another important modality that can provide information on the risk of subsequent acute coronary syndrome (ACS). Although CACS reflects overall coronary atherosclerosis burden and is useful for predicting the risk of CAD, lesion-specific coronary plaque burden and high-risk plaque features, major determinants for ACS risk, can be only assessed by CCTA [191, 192]. However, the analysis of coronary plaque volume and features requires a high level of human expertise and time-consuming protocols, even compared to that needed for CACS measurement. Recent advances in AI have enabled more rapid and accurate assessment of plaque volume and characteristics. Lin et al. used DL model with the hierarchical convolutional long short-term memory network to segment the coronary arteries and showed that this DL-based plaque volume quantification was comparable to that measure by intravascular ultrasound (IVUS), a well-established reference standard, with a shorter analysis time (5.65 s versus 25.66 min for experts) [193]. Araki et al. utilized IVUS in a framework combining SVM and PCA to achieve AUC 0.98 in the risk assessment for CAD [194]. Al'Aref et al. applied XGBoost to train and tune using 10-fold stratified cross-validation on CCTA images and found that this technique could be useful for identifying high-risk plaque features [195]. Han et al. joined the clinical characteristics, biomarkers, and

CCTA-derived variables toward better identify rapid coronary plaque progression in high-risk CAD patients [196]. Li et al. implemented a combined reinforcement multitask progressive time-series networks model using patients' basic patient information with family history, blood biochemical indicators, echocardiography reports, and coronary angiography data on different time to predict the occlusion degree of eight coronary arteries [197]. Combined with traditional risk factors, biomarkers or plaque features, CCTA metrics could aid the risk stratification for CAD risk prediction [198, 199].

Detection of vascular inflammation and novel therapies targeting inflammation have become promising fields of CAD research [200–203]. Peri-coronary and epicardial adipose tissue have attracted growing interest, because these imaging markers can reflect the inflammation. However, their measurement is not considered suitable for clinical practice due to the need of a tedious manual process. A recent study demonstrated that DL model allows fully automated quantification of epicardial adipose tissue with a comparable accuracy and a shorter analysis time (1.57 s versus 15 min for experts) [204]. Furthermore, peri-coronary adipose tissue CT attenuation is the current state-of-the-art method to assess coronary specific inflammation. However, this technique does not account for the complex spatial relationship among voxels. Recent studies suggest that the CT-based radiomics coupled with AI improve the discrimination of MI or the prediction of cardiac risk beyond CT attenuation-based model [205, 206].

## Nuclear Imaging

Current cardiac nuclear imaging is dominated by myocardial perfusion and viability assessment using the flagship techniques of single photon emission computed tomography (SPECT) and positron emission tomography (PET) [207] and its role in the primary prevention of CAD has been limited so far [208, 209]. However, at least two scenarios provide perspectives on cardiac nuclear imaging, expanding their use in primary prevention with the assistance of AI. First, considering that the usage of biomarkers for risk stratification in primary prevention can vary according to their predictive value [210], which can be improved by the implementation of AI, there are still chances that myocardial perfusion imaging (MPI) can be used for primary prevention in the subclinical stage of CAD. Two recent studies highlight this possibility [211]. The first study comparing quantitative versus visual MPI in subtle perfusion defects proved that total perfusion deficit quantified automatically allowed more precise risk stratification [212]. The following study described a DL model significantly surpassing the diagnostic accuracy of standard quantitative analysis and visual reading for MPI [213]. Another possibility of future use of cardiac nuclear imaging for primary prevention of CAD under the assistance of AI can be related to the prediction of high-risk

atherosclerotic plaques associated with a near-term athero-thrombotic event such as myocardial infarction. Clinical studies suggested the inflammatory activities in atherosclerotic plaques measured by 18F-fluorodeoxyglucose (FDG) PET or microcalcification tracked by 18F-sodium fluoride (NaF) PET combined with CT in atherosclerotic plaques are related to cardiovascular events [214–216]. Though there is limited evidence that this can be applied to the primary prevention of CAD, one recent study developed a ML model incorporating quantitative measures of 18F-NaF PET that successfully predicted the future risk of myocardial infarction in patients with stable CAD [217, 218].

## Retinal Imaging

Though there have been several reports on its correlation to CAD, retinal imaging has not been a conventional risk stratification method for the primary prevention of CAD [219, 220]. However, recent studies using DL algorithms shed light on retinal imaging as a potential tool to predict and stratify the risk of CAD. Trained on data from 284,335 patients, the Google AI team employed deep learning algorithms to accurately predict cardiovascular risk factors like age, gender, smoking status, systolic blood pressure, and major cardiac events such as heart attacks from retinal images [221]. In a study using DL algorithms trained on 216,152 retinal images, researchers investigated the algorithm's ability to predict CAC scores and stratify cardiovascular disease risk. The results showed that the DL method, based on retinal images, could predict CAC as determined by CT scans with equal effectiveness in anticipating cardiovascular events [222]. Most recently, after training DL models with retinal and cardiovascular magnetic resonance (CMR) images together, the researchers showed that their algorithm could predict not only the mass or volume of the heart but also future myocardial infarction just using the retinal images and demographic data [223]. Since retinal scans are comparatively cheap and routinely used in many optician practices, with more validation studies, AI-based retinal images might be pushed up as a new risk stratification tool for primary prevention.

## Enhancing CAD Prediction Through AI-Enabled Integration of Personal Health Data and Large Language Models

Increased accessibility of personal EHRs and other digital health data sources provides a rich substrate from which to generate ML-based risk assessment models [224, 225]. Emerging nationwide biobanks have expedited the implementation of these models in care delivery [225, 226]. AI is well suited to parse the sparse yet high dimensional data in EHRs [227]. Recent efforts have begun to demonstrate how genetic risk can be systematically integrated more directly with a wider spectrum of relevant risk factors for risk assessment [228, 229]. ML and NLP approaches have been applied for CVD prediction through parsing structured or unstructured medical big data [230–232]. More specifically, language models — essentially, pre-trained models that can be fine-tuned for various natural language tasks, each of which previously required individual network models — have revolutionized natural language processing (NLP) in recent years. They have become pervasive in NLP, largely due to the success of the transformer architecture [233] and its high compatibility with massively parallel computing hardware. It is now widely recognized that scaling up language models — in terms of training and model parameters — can enhance both performance and sample efficiency across a variety of downstream NLP tasks [234].

To date, one of the largest language models trained with unstructured EHR data — ClinicalBERT [235] — has been developed to characterize reasons for statin nonuse in a multiethnic, real-world, ASCVD cohort. ClinicalBERT includes 110 million parameters and was trained using 0.5 billion words from the publicly available MIMIC-III dataset. The study revealed that around 40% of ASCVD patients lacked formal statin prescriptions. ClinicalBERT effectively detected statin nonuse and primary reasons for this nonuse from unstructured clinical notes — prevalently, patient-level reasons (such as side effects and personal preferences) and clinical-level reasons (i.e., practices that deviate from established guidelines). By guiding targeted interventions to address statin nonuse, clinical LLMs like ClinicalBERT potentially provide a pathway to address important treatment gaps in cardiovascular medicine.

Even larger clinical LLMs have been developed — for example, GatorTron [236] and Med-PaLM 2 [237]. GatorTron scaled up to a size of 8.9 billion parameters using a corpus with 90 billion words from clinical notes, scientific literature, and general English text. It achieved SOTA performance on 5 clinical NLP tasks at various linguistic levels (clinical named entity recognition (CNER), medical relation extraction (MRE), semantic textual similarity (STS), natural language inference (NLI), and medical question answering (MQA)) when compared with three existing clinical/biomedical LM. Remarkably, GatorTron performed considerably better in the most complex NLP tasks (NLI and MQA) compared with existing smaller clinical LMs (BioBERT and ClinicalBERT). Google's Med-PaLM 2 scored up to 86.5% on USMLE MedQA — comparable to an expert doctor — setting a new state-of-the-art and demonstrating the potential of clinical large LM for advanced applications such as MedQA.

A fascinating property of LLMs is emergence, which results from scale [234]. For instance, GPT-3 [238], which boasts 175 billion parameters compared to GPT-2's 1.5 billion, enables in-context learning, in which the LLM can adapt to a specific downstream task simply by receiving a prompt (a natural language description of the task). Intriguingly, this emergent property was neither explicitly trained for nor initially expected to arise [239]. An important consequence of this aspect is the sociological shift within the NLP community toward general-purpose models, i.e., when scaling enables a few-shot prompt-based general-purpose model to outperform previous SOTA performance held by fine-tuned, task-specific models.

Liévin et al. recently explored the capacity of general-purpose LLMs to reason through complex medical questions [239]. Utilizing a human-aligned version of GPT-3 (Instruct-GPT [240]), they addressed multiple-choice questions from medical exams (USMLE and MedQA) as well as medical research queries (PubMedQA). Their investigation employed various techniques: chain-of-thought (CoT) prompts for step-by-step reasoning, grounding by augmenting the prompt with search results, and few-shot learning by prefacing the question with example question–answer pairs. A medical domain expert reviewed and annotated the model's reasoning for a subset of the USMLE questions. Remarkably, even with the most basic prompting schemes, zero-shot GPT-3 outperformed domain-specific BERT baselines. CoT prompting emerged as a particularly effective strategy. By combining multiple CoTs, they discovered that GPT-3 could achieve unprecedented performance on medical questions. Moreover, CoT prompting rendered the zero-shot GPT-3 predictions interpretable, revealing a good comprehension of the context, correct recall of domain-specific knowledge, and non-trivial reasoning patterns. They also noted that the incorporation of few-shot prompt-based learning further improved performance. Lately, Nori et al. ran similar tests on GPT-4 [241], the state-of-the-art LLM at the time of this writing [242]. Without any specialized prompt engineering, GPT-4 exceeded the passing score for the USMLE exam by more than 20 points and outshined earlier general-purpose models (GPT-3.5) as well as models that have been specifically fine-tuned on medical knowledge.

In applications where safety is paramount, such as healthcare, the efficacy of LLMs hinges on their ability to produce outputs that are both factually accurate and comprehensive. The increased conversational abilities of LLMs like GPT-4 enable new paradigms such as multi-agent LLMs. For example, dialog-enabled resolving agents (DERA) is a simple, interpretable forum for models to communicate feedback and iteratively improve output [243]. Dialog is structured as a discussion between two types of agents: a researcher, who processes information and identifies key problem components, and a decider, who has the authority to synthesize the researcher's information and makes final determinations on the output. DERA was evaluated on three tasks with a clinical focus. In the areas of medical conversation summarization and care plan generation, it demonstrated significant improvement over the baseline GPT-4 performance, as evidenced by both human expert preference assessments and quantitative metrics.

Although impressive, these results are not yet on par with human performance. For example, while chain-of-thought prompting approaches suggest the emergence of reasoning patterns that align reasonably well with human approaches to medical problem-solving, they still expose significant gaps in knowledge and reasoning. Interestingly, only the largest GPT models were capable of answering medical questions in a zero-shot setting. This leads to the speculation that the smaller models cannot hold the intricate factual knowledge needed to address specialized medical queries, and that the ability to reason about medical questions only emerges in the largest models. LLMs are expensive to train and require the development of safeguards before being deployed into real-world systems. Notoriously, LLMs have a propensity to magnify the societal biases inherent in their training data, can fabricate information based on the data encoded in their parameters, and it is possible to extract training data from LLM, with larger models being more likely to memorize training data [244, 245]. Therefore, deploying LLMs into sensitive sectors like healthcare must be undertaken with great caution [246, 247]. Nonetheless, LLM are powerful instruments and therefore hold the potential to transform the field of machine intelligence applied to healthcare and primary prevention (CAD and beyond).

## Current Limitations and Future Considerations

The journey to integrate AI into healthcare demands both foundational investment and a cultural shift [248]. This section examines challenges encompassing data availability, data security and privacy, interpretability, and the pivotal role of adequate representation and reduced bias when deploying AI for CAD prevention [249].

Access to diverse and comprehensive datasets is the bedrock of AI's efficacy in healthcare transformation. However, regulatory constraints and fragmented data storage can impede the development and evaluation of AI models. Even though digital health platforms provide vast datasets ideal for AI, the lack of unified data processing and sharing frameworks necessitates considerable curation efforts. Metadata tagging protocols should be standardized to enhance reliability, comparability, and scalability. Accurately harmonizing data from varied platforms and technologies is formidable but indispensable for creating effective AI models.

Initiatives like interdisciplinary consortiums for AI training, technology interfaces for model validation, and open-source sharing of datasets and computational methods are potential solutions [250, 251]. A synchronized strategy for data recording and storage, compatible with diverse devices, is essential [252].

Protecting data throughout AI model lifecycles is paramount. These models, trained on vast and sometimes sensitive datasets, warrant meticulous care to safeguard patient confidentiality. While strategies like data masking and pseudonymization bolster data privacy during AI development, residual risks of data exposure persist. Leveraging AI within decentralized data architectures that emphasize privacy has been proposed [253–255], though its true merit in alleviating privacy concerns is yet to be validated. Unified efforts from researchers, institutions, and regulatory authorities are vital to foster inclusivity during data collection, ensuring AI models that are both potent and secure, thereby benefiting healthcare management and patient outcomes [256, 257].

The intricacies of AI algorithms, often labeled as "black boxes" due to their opaque decision-making processes, present hurdles in building trust among healthcare providers and patients [258]. Successful AI adoption in healthcare necessitates its alignment with clinical practices and guidelines, promoting an interoperable and sustainable care delivery system [259]. Incorporating AI into clinical workflows demands a holistic approach, involving AI-human collaboration among healthcare professionals, data experts, and specialists [260]. This integration is not just about the technology but also about reshaping the decision-making processes in healthcare. Algorithmic solutions targeting lifestyle modifications and emphasizing transparent, actionable predictive pathways are also emerging to address this quandary [261]. The seamless fusion of AI into existing clinical workflows, ensuring interpretability and adherence to guidelines, is of utmost importance [262].

A commitment to inclusivity in healthcare AI is essential, ensuring data covers a wide spectrum of populations. Bias mitigation is vital to prevent AI from inadvertently intensifying health inequities. Yet, datasets harnessed for AI often lack a balance inclusion of diverse ethnic and cultural communities, undermining a model's broad relevance [263]. To counter this, current efforts are underway to diversify ethnic and ancestral makeup of the participant pool [264]. Among them, digital health innovations present opportunities to include often-overlooked groups in medical research, enhancing the accuracy of AI predictions in disease prevention [226, 265]. Tailoring models regionally and adopting transfer learning methods can help bridge performance gaps across demographics [266]. With the flexibility of AI, it could further facilitate precise bias identification and correction, proving superior to conventional risk assessment methods [267]. Such proactive measures are vital for AI's effective integration into healthcare, promoting health equity [268].

In summary, while there are many challenges to implementing AI in healthcare, there are also promising solutions and opportunities to improve patient outcomes. A coordinated effort is needed to address these challenges and to ensure that AI is used ethically and responsibly in healthcare management [269, 270].

## Declarations

**Conflict of Interest** The authors declare no competing interests.

**Human and Animal Rights and Informed Consent** This article does not contain any studies with human or animal subjects performed by any of the authors.

## References

1. Tsao CW, et al. Heart disease and stroke statistics-2022 update: a report from the American Heart Association. Circulation. 2022;145:E153–639.

2. Timmis A, et al. European Society of Cardiology: cardiovascular disease statistics 2021. Eur Heart J. 2022;43:716–99.

3. Libby P. The vascular biology of atherosclerosis. In: Braunwald's heart disease, 2 vol set: a textbook of cardiovascular medicine 12th edition. Elsevier. 2021. 425–441.

4. Morrow D.A, de Lemos J. Stable ischemic heart disease. In: Braunwald's heart disease, 2 vol set: a textbook of cardiovascular medicine 12th edition. Elsevier. 2021. 739–785.

5. Duncker DJ, Canty Jr JM. Coronary blood flow and myocardioal ischemia. In: Braunwald's heart disease, 2 vol set: a textbook of cardiovascular medicine 12th edition. Elsevier, 2021. 609–635.

6. Poulter N. Coronary heart disease is a multifactorial disease. Am J of Hypertens. 1999;12. https://academic.oup.com/ajh/article/12/S6/92S/106550.

7. Aragam KG, et al. Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. Nat Genet. 2022;54:1803–15.

8. Tcheandjieu C, Zhu X, Hilliard AT, Clarke SL, Napolioni V, Ma S, Lee KM, Fang H, Chen F, Lu Y, Tsao NL, Raghavan S, Koyama S, Gorman BR, Vujkovic M, Klarin D, Levin MG, Sinnott-Armstrong N, Wojcik GL, et al. Large-scale genome-wide association study of coronary artery disease in genetically

diverse populations. Nat Med. 2022;28(8):1679–92. https://doi.org/10.1038/s41591-022-01891-3. Accessed 23 Oct 2023.

9. Odden MC, et al. The impact of the aging population on coronary heart disease in the United States. Am J Med. 2011;124:827–33.

10. Mohebi R, et al. Cardiovascular disease projections in the United States based on the 2020 census estimates. J Am Coll Cardiol. 2022;80:565–78.

11. Dai H, et al. Global, regional, and national burden of ischaemic heart disease and its attributable risk factors, 1990–2017: results from the Global Burden of Disease Study 2017. Eur Heart J Qual Care Clin Outcomes. 2022;8:50–60.

12. Chen PB, et al. Directed remodeling of the mouse gut microbiome inhibits the development of atherosclerosis. Nat Biotechnol. 2020;38:1288–97.

13. Quer G, Arnaout R, Henne M, Arnaout R. Machine learning and the future of cardiovascular care: JACC state-of-the-art review. J Am Coll Cardiol. 2021;77:300–13.

14. Shameer K, Johnson KW, Glicksberg BS, Dudley JT, Sengupta PP. Machine learning in cardiovascular medicine: are we there yet? Heart. 2018;104:1156–64.

15. Sardar P, et al. Impact of artificial intelligence on interventional cardiology: from decision-making aid to advanced interventional procedure assistance. JACC Cardiovasc Interv. 2019;12:1293–303.

16. Cho SY, et al. Pre-existing and machine learning-based models for cardiovascular risk prediction. Sci Rep. 2021;11:8886.

17. Christiansen MK, et al. Polygenic risk score-enhanced risk stratification of coronary artery disease in patients with stable chest pain. Circ Genom Precis Med. 2021;14:E003298.

18. Cuocolo R, Perillo T, de Rosa E, Ugga L, Petretta M. Current applications of big data and machine learning in cardiology. J Geriatr Cardiol. 2019;16:601–7.

19. Muse ED, Chen SF, Torkamani A. Monogenic and polygenic models of coronary artery disease. Curr Cardiol Rep. 2021;23:1–12.

20. Olson RS, La Cava W, Mustahsan Z, Varik A, Moore JH. Data-driven advice for applying machine learning to bioinformatics problems. Pac Symp Biocomput. 2018;0:192–203.

21. Krittanawong C, et al. Deep learning for cardiovascular medicine: a practical primer. Eur Heart J. 2019;40:2058-2069C.

22. Belkin M, Hsu D, Ma S, Mandal S. Reconciling modern machine-learning practice and the classical bias–variance tradeoff. Proc Natl Acad Sci U S A. 2019;116:15849–54.

23. Kunapuli G. Ensemble methods: hype or hallelujah? In: Ensemble methods for machine learning 3–20. Manning Publications. 2023.

24. Sapoval N, et al. Current progress and open challenges for applying deep learning across the biosciences. Nat Commun. 2022;13:1–12.

25. Thirunavukarasu AJ, et al. Large language models in medicine. Nat Med. 2023;29:1930–40.

26. Greener JG, Kandathil SM, Moffat L, Jones DT. A guide to machine learning for biologists. Nat Rev Mol Cell Biol. 2022;23:40–55.

27. Krittanawong C, Zhang HJ, Wang Z, Aydar M, Kitai T. Artificial intelligence in precision cardiovascular medicine. J Am Coll Cardiol. 2017;69:2657–64.

28. Hernandez-Boussard T, Monda KL, Crespo BC, Riskin D. Real world evidence in cardiovascular medicine: ensuring data validity in electronic health record-based studies. J Am Med Inform Assoc. 2019;26:1189–94.

29. Azmi J, et al. A systematic review on machine learning approaches for cardiovascular disease prediction using medical big data. Med Eng Phys. 2022;105:103825.

30. Louridi N, Douzi S, El Ouahidi B. Machine learning-based identification of patients with a cardiovascular defect. J Big Data. 2021;8:133. https://doi.org/10.1186/s40537-021-00524-9.

31. Benjamins JW, Hendriks T, Knuuti J, Juarez-Orozco LE, van der Harst P. A primer in artificial intelligence in cardiovascular medicine. Neth Hear J. 2019;27:392–402.

32. Attia Z, Kapa S, Noseworthy P, Friedman P. Artificial intelligence in cardiovascular medicine. In: Braunwald's heart disease, 2 vol set: a textbook of cardiovascular medicine 12th edition. Elsevier. 2021. 109–116.

33. Rigdon J, Basu S. Machine learning with sparse nutrition data to improve cardiovascular mortality risk prediction in the USA using nationally randomly sampled data. BMJ Open. 2019;9:32703.

34. Petrazzini BO, et al. Coronary risk estimation based on clinical data in electronic health records. J Am Coll Cardiol. 2022;79:1155–66.

35. Shorewala V. Early detection of coronary heart disease using ensemble techniques. Inform Med Unlocked. 2021;26:100655. https://doi.org/10.1016/j.imu.2021.100655, https://www.sciencedirect.com/science/article/pii/S235291482100143X

36. Jung S, Ahn E, Koh SB, Lee SH, Hwang GS. Purine metabolite-based machine learning models for risk prediction, prognosis, and diagnosis of coronary artery disease. Biomed Pharmacother. 2021;139:111621. https://doi.org/10.1016/j.biopha.2021.111621.

37. Zuber V, et al. High-throughput multivariable Mendelian randomization analysis prioritizes apolipoprotein B as key lipid risk factor for coronary artery disease. Int J Epidemiol. 2021;50:893–901.

38. Choi MH, Oh S, Choi JY, Kim JH, Lee SW. A statistical learning framework for predicting left ventricular ejection fraction based on glutathione peroxidase-3 level in ischemic heart disease. Comput Biol Med. 2022;149:105929. https://doi.org/10.1016/j.compbiomed.2022.105929.

39. Commandeur F, et al. Machine learning to predict the long-term risk of myocardial infarction and cardiac death based on clinical risk, coronary calcium, and epicardial adipose tissue: a prospective study. Cardiovasc Res. 2020;116:2216–25.

40. Zhang L, et al. Sleep heart rate variability assists the automatic prediction of long-term cardiovascular outcomes. Sleep Med. 2020;67:217–24.

41. Ottosson F, et al. A plasma lipid signature predicts incident coronary artery disease. Int J Cardiol. 2021;331:249–54.

42. Poss AM, et al. Machine learning reveals serum sphingolipids as cholesterol-independent biomarkers of coronary artery disease. J Clin Investig. 2020;130:1363–76.

43. Tsigalou C, et al. Estimation of low-density lipoprotein cholesterol by machine learning methods. Clin Chim Acta. 2021;517:108–16.

44. Vrbaški M, Doroslovački R, Kupusinac A, et al. Lipid profile prediction based on artificial neural networks. J Ambient Intell Human Comput. 2019. https://doi.org/10.1007/s12652-019-01374-3.

45. Snowden SG, Korosi A, de Rooij SR, Koulman A. Combining lipidomics and machine learning to measure clinical lipids in dried blood spots. Metabolomics. 2020;16(8):83. https://doi.org/10.1007/s11306-020-01703-0.

46. Oh GC, et al. Estimation of low-density lipoprotein cholesterol levels using machine learning. Int J Cardiol. 2022;352:144–9.

47. Kesar A, et al. Actionable absolute risk prediction of atherosclerotic cardiovascular disease based on the UK Biobank. PLoS One. 2022;17:2021.11.24.21266742.

48. Marenberg ME, Risch N, Berkman LF, Floderus B, de Faire U. Genetic susceptibility to death from coronary heart disease in a study of twins. N Engl J Med. 1994;330:1041–6.

49. Zdravkovic S, et al. Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. J Intern Med. 2002;252:247–54.

50. Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. Nat Rev Genet. 2018;19:581–90.

51. Chatterjee N, Shi J, García-Closas M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. Nat Rev Genet. 2016;17:392–406.

52. Wells QS, et al. Polygenic risk score to identify subclinical coronary heart disease risk in young adults. Circ Genom Precis Med. 2021;14:e003341.

53. Riveros-Mckay F, et al. Integrated polygenic tool substantially enhances coronary artery disease prediction. Circ Genom Precis Med. 2021;14:E003304.

54. Isgut M, Sun J, Quyyumi AA, Gibson G. Highly elevated polygenic risk scores are better predictors of myocardial infarction risk early in life than later. Genome Med. 2021;13:1–16.

55. Elliott J, et al. Predictive accuracy of a polygenic risk score-enhanced prediction model vs a clinical risk score for coronary artery disease. JAMA - J Am Med Assoc. 2020;323:636–45.

56. Klarin D, Natarajan P. Clinical utility of polygenic risk scores for coronary artery disease. Nat Rev Cardiol. 2022;19:291–301.

57. Mehta NN. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Circulation: Cardiovascular Genetics. 2011;4:327–329.

58. Deloukas P, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet. 2013;45:25–33.

59. Gusev A, et al. Quantifying missing heritability at known GWAS loci. PLoS Genet. 2013;9:e1003993.

60. Won HH, et al. Disproportionate contributions of select genomic compartments and cell types to genetic risk for coronary artery disease. PLoS Genet. 2015;11:e1005622.

61. Khera AV, et al. Genetic risk, adherence to a healthy lifestyle, and coronary disease. New England J Med. 2016;375:2349–58.

62. Mega JL, et al. Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. The Lancet. 2015;385:2264–71.

63. Natarajan P, et al. Polygenic risk score identifies subgroup with higher burden of atherosclerosis and greater relative benefit from statin therapy in the primary prevention setting. Circulation. 2017;135:2091–101.

64. Damask A, et al. Patients with high genome-wide polygenic risk scores for coronary artery disease may receive greater clinical benefit from alirocumab treatment in the ODYSSEY OUTCOMES trial. Circulation. 2020;624–636. https://doi.org/10.1161/CIRCULATIONAHA.119.044434.

65. Marston NA et al. Predicting benefit from evolocumab therapy in patients with atherosclerotic disease using a genetic risk score. Circulation. 2020;616–623. https://doi.org/10.1161/CIRCULATIONAHA.119.043805.

66. Aragam KG, et al. Limitations of contemporary guidelines for managing patients at high genetic risk of coronary artery disease. J Am Coll Cardiol. 2020;75:2769–80.

67. Sattar N, et al. Statins and risk of incident diabetes: a collaborative meta-analysis of randomised statin trials. Lancet. 2010;375:735–42.

68. Macedo AF, et al. Unintended effects of statins from observational studies in the general population: systematic review and meta-analysis. BMC Med. 2014;12:1–13.

69. Bolli A, di Domenico P, Pastorino R, Busby GB, Bottà G. Risk of coronary artery disease conferred by low-density lipoprotein cholesterol depends on polygenic background. Circulation. 2021;143:1452–4.

70. Abraham G, et al. Genomic risk score offers predictive performance comparable to clinical risk factors for ischaemic stroke. Nat Commun. 2019;10:1–10.

71. Krapohl E, et al. Multi-polygenic score approach to trait prediction. Mol Psychiatry. 2018;23:1368–74.

72. Albiñana C, et al. Multi-PGS enhances polygenic prediction by combining 937 polygenic scores. Nat Commun. 2023;14(1):1–11.

73. Mosley JD, et al. Predictive accuracy of a polygenic risk score compared with a clinical risk score for incident coronary heart disease. JAMA - J Am Med Assoc. 2020;323:627–35.

74. Iribarren C, et al. Clinical utility of multimarker genetic risk scores for prediction of incident coronary heart disease: a cohort study among over 51 thousand individuals of European ancestry. Circ Cardiovasc Genet. 2016;9:531–40.

75. Maamari DJ, et al. Clinical implementation of combined monogenic and polygenic risk disclosure for coronary artery disease. JACC: Adv. 2022;1:100068.

76. Patel AP, et al. A multi-ancestry polygenic risk score improves risk prediction for coronary artery disease. Nat Med. 2023;2023:1–11. https://doi.org/10.1038/s41591-023-02429-x.

77. Naushad SM, et al. Machine learning algorithm-based risk prediction model of coronary artery disease. Mol Biol Rep. 2018;45:901–10.

78. Zhou K, Cai C, He Y, Chen Z. Potential prognostic biomarkers of sudden cardiac death discovered by machine learning. Comput Biol Med. 2022;106154.https://doi.org/10.1016/J.COMPBIOMED.2022.106154.

79. Manduchi E, Le TT, Fu W, Moore JH. Genetic analysis of coronary artery disease using tree-based automated machine learning informed by biology-based feature selection. IEEE/ACM Trans Comput Biol Bioinform. 2022;19:1379–86.

80. Mishra A, et al. Stroke genetics informs drug discovery and risk prediction across ancestries. Nature. 2022. https://doi.org/10.1038/s41586-022-05165-3.

81. Dey KK, et al. Integrative approaches to improve the informativeness of deep learning models for human complex diseases. BioRxiv. 2021;2020.09.08.288563.https://doi.org/10.1101/2020.09.08.288563.

82. Caron B, Luo Y, Rausell A. NCBoost classifies pathogenic noncoding variants in Mendelian diseases through supervised learning on purifying selection signals in humans. Genome Biol. 2019;20(1):32. https://doi.org/10.1186/s13059-019-1634-2.

83. Wang QS, Kelley DR, Ulirsch J, Kanai M, Sadhuka S, Cui R, Albors C, Cheng N, Okada Y, Project BJ, Aguet F, Ardlie KG, DG MA, Finucane HK. Leveraging supervised learning for functionally informed fine-mapping of cis-eQTLs identifies an additional 20,913 putative causal eQTLs. Nat Commun. 2021;12(1):3394. https://doi.org/10.1038/s41467-021-23134-8.

84. Huang S, Ji X, Cho M, Joo J, Moore J. DL-PRS: a novel deep learning approach to polygenic risk scores. Res Sq. 2021;1–14. https://doi.org/10.21203/RS.3.RS-423764/V1.

85. Jiajie Peng A, et al. A deep learning-based genome-wide polygenic risk score for common diseases identifies individuals with risk. MedRxiv. 2021;2021.11.17.21265352. https://doi.org/10.1101/2021.11.17.21265352.

86. Koido M, et al. Prediction of the cell-type-specific transcription of non-coding RNAs from genome sequences via machine learning. Nat Biomed Eng. 2022;7(6):830–44.

87. Turner AW, et al. Single-nucleus chromatin accessibility profiling highlights regulatory mechanisms of coronary artery disease risk. Nat Genet. 2022;54:804–16.

88. Amariuta T, et al. IMPACT: genomic annotation of cell-state-specific regulatory elements inferred from the epigenome of bound transcription factors. Am J Hum Genet. 2019;104:879–95.

89. Amariuta T, et al. Improving the trans-ancestry portability of polygenic risk scores by prioritizing variants in predicted cell-type-specific regulatory elements. Nat Genet. 2020;52:1346–54.

90. Zhang T, et al. GCN-GENE: a novel method for prediction of coronary heart disease-related genes. Comput Biol Med 105918 (2022). https://doi.org/10.1016/j.compbiomed.2022.105918.

91. van Hilten A, et al. GenNet framework: interpretable deep learning for predicting phenotypes from genetic data. Commun Biol. 2021;4:1094.

92. Medvedev A, Mishra Sharma S, Tsatsorin E, Nabieva E, Yarotsky D. Human genotype-to-phenotype predictions: boosting accuracy with nonlinear models. PLoS One. 2022;17:e0273293.

93. Elgart M, Lyons G, Romero-Brufau S, Kurniansyah N, Brody JA, Guo X, Lin HJ, Raffield L, Gao Y, Chen H, de Vries P, Lloyd-Jones DM, Lange LA, Peloso GM, Fornage M, Rotter JI, Rich SS, Morrison AC, Psaty BM, et al. Non-linear machine learning models incorporating SNPs and PRS improve polygenic prediction in diverse human populations. Commun Biol. 2022;5(1):856. https://doi.org/10.1038/s42003-022-03812-z.

94. Lau M, Wigmann C, Kress S, Schikowski T, Schwender H. Evaluation of tree-based statistical learning methods for constructing genetic risk scores. BMC Bioinform. 2022;23:1–30.

95. Kovacic JC. Unraveling the complex genetics of coronary artery disease. J Am Coll Cardiol. 2017;69:837–40.

96. Hartiala JA, Hilser JR, Biswas S, Lusis AJ, Allayee H. Gene-environment interactions for cardiovascular disease. Curr Atheroscler Rep. 2021;23:1–9.

97. Flores AM, et al. Unsupervised learning for automated detection of coronary artery disease subgroups. J Am Heart Assoc. 2021;10:21976.

98. Agrawal S, Klarqvist MDR, Emdin C, Patel AP, Paranjpe MD, Ellinor PT, Philippakis A, Ng K, Batra P, Khera AV. Selection of 51 predictors from 13,782 candidate multimodal features using machine learning improves coronary artery disease prediction. Patterns (N Y). 2021;2(12):100364. https://doi.org/10.1016/j.patter.2021.100364.

99. Xu Y, et al. A machine learning model for disease risk prediction by integrating genetic and non-genetic factors. Proc - 2022 IEEE Int Conf Bioinform Biomed, BIBM. 2022. 868–871. https://doi.org/10.1109/BIBM55620.2022.9994925.

100. Dogan MV, Grumbach IM, Michaelson JJ, Philibert RA. Integrated genetic and epigenetic prediction of coronary heart disease in the Framingham Heart Study. PLoS ONE. 2018;13:e0190549.

101. Forrest IS, et al. Machine learning-based marker for coronary artery disease: derivation and validation in two longitudinal cohorts. Lancet. 2023;401:215–25.

102. Nam Y, et al. netCRS: network-based comorbidity risk score for prediction of myocardial infarction using biobank-scaled PheWAS data. Pac Symp Biocomput. 2022;27:325–36.

103. Steinfeldt J, et al. Neural network-based integration of polygenic and clinical information: development and validation of a prediction model for 10-year risk of major adverse cardiac events in the UK Biobank cohort. Lancet Digit Health. 2022;4:e84–94.

104. Westerlund AM, Hawe JS, Heinig M, Schunkert H. Risk prediction of cardiovascular events by exploration of molecular data with explainable artificial intelligence. Int J Mol Sci. 2021;22:10291.

105. Muse ED, Topol EJ. Digital orthodoxy of human data collection. The Lancet. 2019;394:556.

106. Muse ED, et al. Impact of polygenic risk communication: an observational mobile application-based coronary artery disease study. NPJ Digit Med. 2022;5.

107. Linder JE, et al. Returning integrated genomic risk and clinical recommendations: the eMERGE study. Genet Med. 2023;25:100006.

108. Chen B, et al. Harnessing big 'omics' data and AI for drug discovery in hepatocellular carcinoma. Nat Rev Gastroenterol Hepatol. 2020;17(4):238–51.

109. Odenkirk MT, Reif DM, Baker ES. Multiomic big data analysis challenges: increasing confidence in the interpretation of artificial intelligence assessments. Anal Chem. 2021;93:7763–73.

110. Reel PS, Reel S, Pearson E, Trucco E, Jefferson E. Using machine learning approaches for multi-omics data analysis: a review. Biotechnol Adv. 2021;49:107739.

111. Muse ED, et al. A Whole Blood Molecular Signature for Acute Myocardial Infarction. Sci Rep. 2017;7(1):1–9.

112. Kang L, et al. Uncovering potential diagnostic biomarkers of acute myocardial infarction based on machine learning and analyzing its relationship with immune cells. BMC Cardiovasc Disord. 2023;23:1–12.

113. Yifan C, Jianfeng S, Jun P. Development and validation of a random forest diagnostic model of acute myocardial infarction based on ferroptosis-related genes in circulating endothelial cells. Front Cardiovasc Med. 2021;8:444.

114. Damani S, Bacconi A, Libiger O, Chourasia AH, Serry R, Gollapudi R, Goldberg R, Rapeport K, Haaser S, Topol S, Knowlton S, Bethel K, Kuhn P, Wood M, Carragher B, Schork NJ, Jiang J, Rao C, Connelly M, et al. Characterization of circulating endothelial cells in acute myocardial infarction. Sci Transl Med. 2012;4(126):126ra33. https://doi.org/10.1126/scitranslmed.3003451.

115. Zhao X, et al. A deep learning model for early risk prediction of heart failure with preserved ejection fraction by DNA methylation profiles combined with clinical features. Clin Epigenetics. 2022;14:1–15.

116. Westerman K, et al. Epigenomic assessment of cardiovascular disease risk and interactions with traditional risk metrics. J Am Heart Assoc. 2020;9:15299.

117. Williams SA, et al. A proteomic surrogate for cardiovascular outcomes that is sensitive to multiple mechanisms of change in risk. Sci Transl Med. 2022;14:9625.

118. Unterhuber M, et al. Proteomics-enabled deep learning machine algorithms can enhance prediction of mortality. J Am Coll Cardiol. 2021;78:1621–31.

119. Khera R, et al. Use of machine learning models to predict death after acute myocardial infarction. JAMA Cardiol. 2021;6:633–41.

120. Hoogeveen RM, et al. Improved cardiovascular risk prediction using targeted plasma proteomics in primary prevention. Eur Heart J. 2020;41:3998–4007.

121. Nurmohamed NS, et al. Targeted proteomics improves cardiovascular risk prediction in secondary prevention. Eur Heart J. 2022;43:1569–77.

122. Sze SK. Artificially intelligent proteomics improves cardiovascular risk assessment. EBioMedicine. 2019;40:23–4.

123. Aryal S, Alimadadi A, Manandhar I, Joe B, Cheng X. Machine learning strategy for gut microbiome-based diagnostic screening of cardiovascular disease. Hypertension. 2020;76:1555–62.

124. Jie Z, et al. The gut microbiome in atherosclerotic cardiovascular disease. Nat Commun. 2017;8:1–12.

125. Jonsson AL, Bäckhed F. Role of gut microbiota in atherosclerosis. Nat Rev Cardiol. 2017;14:79–87.

126. Masenga SK, et al. Recent advances in modulation of cardiovascular diseases by the gut microbiota. J Hum Hypertens. 2022;36:952–9.

127. Zhu Q, et al. Dysbiosis signatures of gut microbiota in coronary artery disease. Physiol Genomics. 2018;50:893–903.

128. Bayoumy K, et al. Smart wearable devices in cardiovascular care: where we are and how to move forward. Nat Rev Cardiol. 2021;18:581–99.

129. Krittanawong C, et al. Integration of novel monitoring devices with machine learning technology for scalable cardiovascular management. Nat Rev Cardiol. 2021;18:75–91.

130. Engelhard MM, Oliver JA, McClernon FJ. Digital envirotyping: quantifying environmental determinants of health and

behavior. NPJ Digit Med. 2020;3:36. https://doi.org/10.1038/s41746-020-0245-3.

131. Beckers J, Wurst W, de Angelis MH. Towards better mouse models: enhanced genotypes, systemic phenotyping and envirotype modelling. Nat Rev Genet. 2009;10:371–80.

132. Zhang Y, Hu Y, Jiang N, Yetisen AK. Wearable artificial intelligence biosensor networks. Biosens Bioelectron. 2023;219:114825.

133. Zhang M, Cui X, Li N. Smartphone-based mobile biosensors for the point-of-care testing of human metabolites. Mater Today Bio. 2022;14:100254.

134. Xintarakou A, Sousonis V, Asvestas D, Vardas PE, Tzeis S. Remote cardiac rhythm monitoring in the era of smart wearables: present assets and future perspectives. Front Cardiovasc Med. 2022;9:853614.

135. Friedman P, Vardas PE, Asselbergs FW, van Smeden M. The year in cardiovascular medicine 2021: digital health and innovation. Eur Heart J. 2022;43:271–9.

136. Kumar S, et al. Wearables in cardiovascular disease. J Cardiovasc Transl Res. 2023;16:557–68.

137. Burman A, Titus J, Gbadebo D, Burman M. ECGDetect: detecting ischemia via deep learning. arXiv preprint arXiv:2009.13232. 2020. https://doi.org/10.48550/arXiv.2009.13232.

138. al Hinai G, Jammoul S, Vajihi Z, Afilalo J. Deep learning analysis of resting electrocardiograms for the detection of myocardial dysfunction, hypertrophy, and ischaemia: a systematic review. European Heart J - Digit Health. 2021;2:416–23.

139. Jansi Rani SV, et al. Smart wearable model for predicting heart disease using machine learning: wearable to predict heart risk. J Ambient Intell Humaniz Comput. 2022;13:4321–32.

140. Sana F, et al. Wearable devices for ambulatory cardiac monitoring: JACC state-of-the-art review. J Am Coll Cardiol. 2020;75:1582–92.

141. Attia ZI, et al. Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. Nat Med. 2019;25:70–4.

142. He Z, et al. Hemorrhagic risk prediction in coronary artery disease patients based on photoplethysmography and machine learning. Sci Reports. 2022;12(1):1–12.

143 Siontis KC, Noseworthy PA, Attia ZI, Friedman PA. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. Nat Rev Cardiol. 2021;18(7):465–78.

144. Reiss A, Indlekofer I, Schmidt P, Van Laerhoven K. Deep PPG: large-scale heart rate estimation with convolutional neural networks. Sensors (Basel). 2019;19(14):3079. https://doi.org/10.3390/s19143079.

145. Sun X, Zhou L, Chang S, Liu Z. Using CNN and HHT to predict blood pressure level based on photoplethysmography and its derivatives. Biosensors (Basel). 2021;11(4):120. https://doi.org/10.3390/bios11040120.

146. Gupta K, Bajaj V, Ansari IA, Rajendra Acharya U. Hyp-Net: automated detection of hypertension using deep convolutional neural network and Gabor transform techniques with ballistocardiogram signals. Biocybern Biomed Eng. 2022;42:784–96.

147. Yang C, Ojha BD, Aranoff ND, Green P, Tavassolian N. Classification of aortic stenosis using conventional machine learning and deep learning methods based on multi-dimensional cardio-mechanical signals. Sci Rep. 2020;10:1–11.

148. Sivasangari A, et al. An artificial intelligence mediated integrated wearable device for diagnosis of cardio through remote monitoring. Cogn Syst Signal Proc Image Proc. 2021;319–335. https://doi.org/10.1016/B978-0-12-824410-4.00008-8.

149. Sahoo PK, Mohapatra S, Thakkar HK. Artificial intelligence assisted cardiac signal analysis for heart disease prediction. EAI/Springer Innov Commun Comput. 2023;337–372. https://doi.org/10.1007/978-3-031-15816-2_18.

150. Khan MU, et al. Artificial neural network-based cardiovascular disease prediction using spectral features. Comput Electr Eng. 2022;101:108094.

151. Yang W, et al. Diagnosis of cardiac abnormalities based on phonocardiogram using a novel fuzzy matching feature extraction method. BMC Med Inform Decis Mak. 2022;22:1–13.

152. Hannun AY, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat Med. 2019;25:65–9.

153 Kwon JM, et al. Artificial intelligence-enhanced smartwatch ECG for heart failure-reduced ejection fraction detection by generating 12-lead ECG. Diagnostics. 2022;12:654.

154. Vasudeva ST, et al. Development of a convolutional neural network model to predict coronary artery disease based on single-lead and twelve-lead ECG signals. Appl Sci. 2022;12:7711.

155. Mora S, Cook N, Buring JE, Ridker PM, Lee IM. Physical activity and reduced risk of cardiovascular events: potential mediating mechanisms. Circulation. 2007;116:2110–8.

156. Wallace ML, et al. Multidimensional sleep and mortality in older adults: a machine-learning comparison with other risk factors. J Gerontol - Ser A Biol Sci Med Sci. 2019;74:1903–9.

157. Lee S, et al. Sleep health composites are associated with the risk of heart disease across sex and race. Sci Rep. 2022;12:1–11.

158. Nguyen S, et al. Accelerometer-derived daily life movement classified by machine learning and incidence of cardiovascular disease in older women: the OPACH study. J Am Heart Assoc. 2022;11:23433.

159. Paluch AE, et al. Steps per day and all-cause mortality in middle-aged adults in the coronary artery risk development in young adults study. JAMA Netw Open. 2021;4:e2124516–e2124516.

160. Albalak G, et al. Setting your clock: associations between timing of objective physical activity and cardiovascular disease risk in the general population. Eur J Prev Cardiol. 2023;30:232–40.

161. Dempsey PC, et al. Physical activity volume, intensity, and incident cardiovascular disease. Eur Heart J. 2022;43:4789–800.

162. Huang W, et al. Application of ensemble machine learning algorithms on lifestyle factors and wearables for cardiovascular risk prediction. Sci Rep. 2022;12:1033.

163. Prabhu G, O'connor NE, Moran K. Recognition and repetition counting for local muscular endurance exercises in exercise-based rehabilitation: a comparative study using artificial intelligence models. Sensors (Switzerland). 2020;20:1–29.

164. Su J, Zhang Y, Ke QQ, Su JK, Yang QH. Mobilizing artificial intelligence to cardiac telerehabilitation. Rev Cardiovasc Med. 2022;23:45.

165. Batalik L, et al. Long-term exercise effects after cardiac telerehabilitation in patients with coronary artery disease: 1-year follow-up results of the randomized study. Eur J Phys Rehabil Med. 2021;57:807–14.

166. Antoniou V, et al. Efficacy, efficiency and safety of a cardiac telerehabilitation programme using wearable sensors in patients with coronary heart disease: the TELEWEAR-CR study protocol. BMJ Open. 2022;12:e059945.

167. Yang P, Wei G, Liu A, Huo F, Zhang Z. A review of sampling, energy supply and intelligent monitoring for long-term sweat sensors. npj Flex Electron. 2022;6:1–13.

168. Vincent Paul SM, Balasubramaniam S, Panchatcharam P, Malarvizhi Kumar P, Mubarakali A. Intelligent framework for prediction of heart disease using deep learning. Arab J Sci Eng. 2022;47:2159–69.

169. Dunn J, et al. Wearable sensors enable personalized predictions of clinical laboratory measurements. Nat Med. 2021;27:1105–12.

170. Anusha AS, Preejith SP, Akl TJ, Sivaprakasam M. Electrodermal activity based autonomic sleep staging using wrist wearable. Biomed Signal Process Control. 2022;75:103562.

171. Tozlu BH, Şimşek C, Aydemir O, Karavelioglu Y. A high performance electronic nose system for the recognition of myocardial infarction and coronary artery diseases. Biomed Signal Process Control. 2021;64:102247.

172. Sohail A, Ashiq U. Quantum inspired improved AI computing for the sensors of cardiac mechano-biology. Sensors International. 2023;4:100212.

173. Vashistha R, Dangi AK, Kumar A, Chhabra D, Shukla P. Futuristic biosensors for cardiac health care: an artificial intelligence approach. 3 Biotech. 2018;8:1–11.

174. Powell-Wiley TM, et al. Social determinants of cardiovascular disease. Circ Res. 2022;130:782–99.

175. Li J, Li X, Zhang S, Snyder M. Gene-environment interaction in the era of precision medicine. Cell. 2019;177:38–44.

176. Zhou W, Chan YE, Foo CS, Zhang J, Teo JX, Davila S, Huang W, Yap J, Cook S, Tan P, Chin CW, Yeo KK, Lim WK, Krishnaswamy P. High-resolution digital phenotypes from consumer wearables and their applications in machine learning of cardiometabolic risk markers: Cohort study. J Med Internet Res. 2022;24(7):e34669. https://doi.org/10.2196/34669. Accessed 23 Oct 2023.

177. Marabita F, et al. Multiomics and digital monitoring during lifestyle changes reveal independent dimensions of human biology and health. Cell Syst. 2022;13:241-255.e7.

178. Jalali S, et al. Long-term exposure to PM2.5 and cardiovascular disease incidence and mortality in an Eastern Mediterranean country: findings based on a 15-year cohort study. Environ Health. 2021;20:1–16.

179. Mirowsky JE, Carraway MS, Dhingra R, Tong H, Neas L, Diaz-Sanchez D, Cascio WE, Case M, Crooks JL, Hauser ER, Dowdy ZE, Kraus WE, Devlin RB. Exposures to low-levels of fine particulate matter are associated with acute changes in heart rate variability, cardiac repolarization, and circulating blood lipids in coronary artery disease patients. Environ Res. 2022;214(Pt 1):113768. https://doi.org/10.1016/j.envres.2022.113768.

180. Li X, et al. Digital health: tracking physiomes and activity using wearable biosensors reveals useful health-related information. PLoS Biol. 2017;15:e2001402.

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

182. Kaptoge S, et al. World Health Organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions. Lancet Glob Health. 2019;7:e1332–45.

183. Patnode CD, Redmond N, Iacocca MO, Henninger M. Behavioral counseling interventions to promote a healthy diet and physical activity for cardiovascular disease prevention in adults without known cardiovascular disease risk factors: updated evidence report and systematic review for the US Preventive Services Task Force. JAMA. 2022;328:375–88.

184. Detrano R, Guerci AD, Carr JJ, Bild DE, Burke G, Folsom AR, Liu K, Shea S, Szklo M, Bluemke DA, O'Leary DH, Tracy R, Watson K, Wong ND, Kronmal RA. Coronary calcium as a predictor of coronary events in four racial or ethnic groups. N Engl J Med. 2008;358(13):1336–45. https://doi.org/10.1056/NEJMoa072100. Accessed 23 Oct 2023.

185. Yeboah, J. et al. Comparison of novel risk markers for improvement in cardiovascular risk assessment in intermediate-risk individuals. Jama. 2012;308. http://www.mesa-nhlbi.org.

186. Agarwal S, et al. Coronary calcium score predicts cardiovascular mortality in diabetes: diabetes heart study. Diabetes Care. 2013;36:972–7.

187. Sandstedt M, et al. Evaluation of an AI-based, automatic coronary artery calcium scoring software. Eur Radiol. 2020;30:1671–8.

188. van Velzen SGM, et al. Deep learning for automatic calcium scoring in CT: validation using multiple cardiac CT and chest CT protocols. Radiology. 2020;295:66–79.

189. Eng D, Chute C, Khandwala N, Rajpurkar P, Long J, Shleifer S, Khalaf MH, Sandhu AT, Rodriguez F, Maron DJ, Seyyedi S, Marin D, Golub I, Budoff M, Kitamura F, Takahashi MS, Filice RW, Shah R, Mongan J, et al. Automated coronary calcium scoring using deep learning with multicenter external validation. NPJ Digit Med. 2021;4(1):88. https://doi.org/10.1038/s41746-021-00460-1.

190. Kelly-Arnold A, et al. Revised microcalcification hypothesis for fibrous cap rupture in human coronary arteries. Proc Natl Acad Sci U S A. 2013;110:10741–6.

191. Chang HJ, et al. Coronary atherosclerotic precursors of acute coronary syndromes. J Am Coll Cardiol. 2018;71:2511–22.

192. Andreini D, et al. Coronary plaque features on CTA can identify patients at increased risk of cardiovascular events. JACC Cardiovasc Imaging. 2020;13:1704–17.

193. Lin A, 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.

194. Araki T, et al. PCA-based polling strategy in machine learning framework for coronary artery disease risk assessment in intravascular ultrasound: a link between carotid and coronary grayscale plaque morphology. Comput Methods Programs Biomed. 2016;128:137–58.

195. Al'Aref SJ, et al. A boosted ensemble algorithm for determination of plaque stability in high-risk patients on coronary CTA. JACC Cardiovasc Imaging. 2020;13:2162–73.

196. Han D, Kolli KK, Al'Aref SJ, Baskaran L, van Rosendael AR, Gransar H, Andreini D, Budoff MJ, Cademartiri F, Chinnaiyan K, Choi JH, Conte E, Marques H, de Araújo GP, Gottlieb I, Hadamitzky M, Leipsic JA, Maffei E, Pontone G, et al. machine learning framework to identify individuals at risk of rapid progression of coronary atherosclerosis: from the PARADIGM registry. J Am Heart Assoc. 2020;9(5):e013958. https://doi.org/10.1161/JAHA.119.013958.

197. Li W, Zuo M, Zhao H, Xu Q, Chen D. Prediction of coronary heart disease based on combined reinforcement multitask progressive time-series networks. Methods. 2022;198:96–106.

198. Li D, et al. Machine learning-aided risk stratification system for the prediction of coronary artery disease. Int J Cardiol. 2021;326:30–4.

199. Yang S, et al. CT angiographic and plaque predictors of functionally significant coronary disease and outcome using machine learning. JACC Cardiovasc Imaging. 2021;14:629–41.

200. Antonopoulos AS, et al. Detecting human coronary inflammation by imaging perivascular fat. Sci Trans Med. 2017;9. https://www.science.org.

201. Packer M. Epicardial adipose tissue may mediate deleterious effects of obesity and inflammation on the myocardium. J Am Coll Cardiol. 2018;71:2360–72.

202. Ridker PM, et al. Antiinflammatory therapy with canakinumab for atherosclerotic disease. N Engl J Med. 2017;377:1119–31.

203. Nidorf SM, et al. Colchicine in patients with chronic coronary disease. N Engl J Med. 2020;383:1838–47.

204. Commandeur F, et al. Fully automated CT quantification of epicardial adipose tissue by deep learning: a multicenter study. Radiol Artif Intell. 2019;1.

205. Lin A, et al. Myocardial infarction associates with a distinct pericoronary adipose tissue radiomic phenotype: a prospective case-control study. JACC Cardiovasc Imaging. 2020;13:2371–83.

206. Oikonomou EK, et al. A novel machine learning-derived radiotranscriptomic signature of perivascular fat improves cardiac

risk prediction using coronary CT angiography. Eur Heart J. 2019;40:3529–43.

207. Juarez-Orozco LE, et al. Artificial intelligence to improve risk prediction with nuclear cardiac studies. Curr Cardiol Rep. 2022;24:307–16.

208. Beller GA, Zaret BL. Contributions of nuclear cardiology to diagnosis and prognosis of patients with coronary artery disease. Circulation. 2000;101(12):1465–78. https://doi.org/10.1161/01.cir.101.12.1465. Accessed 23 Oct 2023.

209. Berman DS, et al. Roles of nuclear cardiology, cardiac computed tomography, and cardiac magnetic resonance: assessment of patients with suspected coronary artery disease. J Nucl Med. 2006;47:74–82.

210. Wang TJ. Assessing the role of circulating, genetic, and imaging biomarkers in cardiovascular risk prediction. Circulation. 2011;123:551–65.

211. Otaki Y, Miller RJH, Slomka PJ. The application of artificial intelligence in nuclear cardiology. Ann Nucl Med. 2022;36:111–22.

212. Otaki Y, et al. 5-Year prognostic value of quantitative versus visual MPI in subtle perfusion defects: results from REFINE SPECT. JACC Cardiovasc Imaging. 2020;13:774–85.

213. Otaki Y, et al. Clinical deployment of explainable artificial intelligence of SPECT for diagnosis of coronary artery disease. JACC Cardiovasc Imaging. 2022;15:1091–102.

214. Figueroa AL, et al. Measurement of arterial activity on routine FDG PET/CT images improves prediction of risk of future CV events. JACC Cardiovasc Imaging. 2013;6:1250–9.

215. Rominger A, et al. 18F-FDG PET/CT identifies patients at risk for future vascular events in an otherwise asymptomatic cohort with neoplastic disease. J Nucl Med. 2009;50:1611–20.

216. Joshi NV, et al. 18F-fluoride positron emission tomography for identification of ruptured and high-risk coronary atherosclerotic plaques: a prospective clinical trial. Lancet. 2014;383:705–13.

217. Kwiecinski J, et al. Machine learning with 18F-sodium fluoride pet and quantitative plaque analysis on CT angiography for the future risk of myocardial infarction. J Nucl Med. 2022;63:158–65.

218. Popescu C, et al. PET-based artificial intelligence applications in cardiac nuclear medicine. Swiss Med Wkly. 2022;152.

219. Witt N, et al. Abnormalities of retinal microvascular structure and risk of mortality from ischemic heart disease and stroke. Hypertension. 2006;47:975–81.

220. Wang SB, et al. A spectrum of retinal vasculature measures and coronary artery disease. Atherosclerosis. 2018;268:215–24.

221. Poplin R, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. Nat Biomed Eng. 2018;2:158–64.

222. Rim TH, et al. Deep-learning-based cardiovascular risk stratification using coronary artery calcium scores predicted from retinal photographs. Lancet Digit Health. 2021;3:e306–16.

223. Diaz-Pinto A, et al. Predicting myocardial infarction through retinal scans and minimal personal information. Nat Mach Intell. 2022;4:55–61.

224. Dai H, et al. Big data in cardiology: state-of-art and future prospects. Front Cardiovasc Med. 2022;9:844296.

225. Abul-Husn NS, Kenny EE. Personalized medicine and the power of electronic health records. Cell. 2019;177:58–69.

226. Zhou W, et al. Global Biobank Meta-analysis Initiative: powering genetic discovery across human disease. Cell Genomics. 2022;2:2021.11.19.21266436.

227. Hossain E, et al. Natural language processing in electronic health records in relation to healthcare decision-making: a systematic review. Comput Biol Med. 2023;155:106649.

228. Dennis JK, et al. Clinical laboratory test-wide association scan of polygenic scores identifies biomarkers of complex disease. Genome Med. 2021;13:1–16.

229. Dolezalova N, et al. Development of an accessible 10-year Digital CArdioVAscular (DiCAVA) risk assessment: a UK Biobank study. European Heart J - Digit Health. 2021;2:528–38.

230. Alaa AM, Bolton T, Angelantonio ED, Rudd JHF, van der Schaar M. Cardiovascular disease risk prediction using automated machine learning: a prospective study of 423,604 UK Biobank participants. PLoS ONE. 2019;14:e0213653.

231. Ward A, Sarraju A, Chung S, Li J, Harrington R, Heidenreich P, Palaniappan L, Scheinker D, Rodriguez F. Machine learning and atherosclerotic cardiovascular disease risk prediction in a multi-ethnic population. NPJ Digit Med. 2020;3:125. https://doi.org/10.1038/s41746-020-00331-1.

232. Banda JM, Sarraju A, Abbasi F, Parizo J, Pariani M, Ison H, Briskin E, Wand H, Dubois S, Jung K, Myers SA, Rader DJ, Leader JB, Murray MF, Myers KD, Wilemon K, Shah NH, Knowles JW. Finding missed cases of familial hypercholesterolemia in health systems using machine learning. NPJ Digit Med. 2019;2:23. https://doi.org/10.1038/s41746-019-0101-5.

233. Vaswani A, et al. Attention is all you need. Adv Neural Inf Process Syst. 2017; 2017-Decem:5999–6009.

234. Wei J, Tay Y, Bommasani R, Raffel C, Zoph B, Borgeaud S, Yogatama D, Bosma M, Zhou D, Metzler D, Chi EH. Emergent abilities of large language models. arXiv preprint arXiv:2206.07682. 2022. https://doi.org/10.48550/arXiv.2206.07682.

235. Sarraju A, Coquet J, Zammit A, Chan A, Ngo S, Hernandez-Boussard T, Rodriguez F. Using deep learning-based natural language processing to identify reasons for statin nonuse in patients with atherosclerotic cardiovascular disease. Commun Med (Lond). 2022;2:88. https://doi.org/10.1038/s43856-022-00157-w.

236. Yang X, et al. A large language model for electronic health records. npj Digital Med. 2022;5(1):1–9.

237. Singhal K, et al. Towards expert-level medical question answering with large language models. ArXiv. 2023. https://doi.org/10.48550/arXiv.2305.09617.

238. Brown TB, et al. Language models are few-shot learners. Adv Neural Inf Process Syst. 2020;2020-Decem.

239. Liévin V, Hother CE, Winther O. Can large language models reason about medical questions? ArXiv. 2022. https://doi.org/10.48550/arXiv.2207.08143.

240. Ouyang L, et al. Training language models to follow instructions with human feedback. ArXiv. 2022. https://doi.org/10.48550/arxiv.2203.02155.

241. Nori H, King N, McKinney SM, Carignan D, Horvitz E. Capabilities of GPT-4 on medical challenge problems. ArXiv. 2023. https://doi.org/10.48550/arXiv.2303.13375.

242. Lee P, Bubeck S, Petro J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. N Engl J Med. 2023;388:1233–9.

243. Nair V, Schumacher E, Tso G, Kannan A. DERA: enhancing large language model completions with dialog-enabled resolving agents. ArXiv. 2023. https://doi.org/10.48550/arXiv.2303.17071.

244. Carlini N, et al. Quantifying memorization across neural language models. ArXiv. 2022. https://doi.org/10.48550/arXiv.2202.07646.

245. Meskó B, Topol EJ. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. NPJ Digit Med. 2023;6:1–6.

246. Korngiebel DM, Mooney SD. Considering the possibilities and pitfalls of Generative Pre-trained Transformer 3 (GPT-3) in healthcare delivery. NPJ Digit Med. 2021;4(1):93. https://doi.org/10.1038/s41746-021-00464-x.

247. Sezgin E, Sirrianni J, Linwood SL. Operationalizing and implementing pretrained, large artificial intelligence linguistic models in the US Health Care System: outlook of generative pretrained

transformer 3 (GPT-3) as a service model. JMIR Med Inform. 2022;10(2):e32875. https://doi.org/10.2196/32875.

248. Dias R, Torkamani A. Artificial intelligence in clinical and genomic diagnostics. Genome Med. 2019;11:1–12.

249. He J, et al. The practical implementation of artificial intelligence technologies in medicine. Nat Med. 2019;25:30–6.

250. Brown SA, et al. Establishing an interdisciplinary research team for cardio-oncology artificial intelligence informatics precision and health equity. Am Heart J Plus: Cardiol Res Pract. 2022;13:100094.

251. Vishwanatha JK, et al. Community perspectives on AI/ML and health equity: AIM-AHEAD nationwide stakeholder listening sessions. PLOS Digital Health. 2023;2:e0000288.

252. Acosta JN, Falcone GJ, Rajpurkar P, Topol EJ. Multimodal biomedical AI. Nat Med. 2022;28(9):1773–84.

253. Puri V, Kataria A, Sharma V. Artificial intelligence-powered decentralized framework for Internet of Things in Healthcare 4.0. Trans Emerg Telecommun Technol. 2021;e4245. https://doi.org/10.1002/ett.4245.

254. Warnat-Herresthal S, et al. Swarm learning for decentralized and confidential clinical machine learning. Nature. 2021;594:265–70.

255. De Brouwer W, Patel CJ, Manrai AK, Rodriguez-Chavez IR, Shah NR. Empowering clinical research in a decentralized world. NPJ Digit Med. 2021;4:1–5.

256. Ghassemi M, et al. Practical guidance on artificial intelligence for health-care data. Lancet Digit Health. 2019;1:e157–9.

257. Chen IY, et al. Ethical machine learning in healthcare. Annu Rev Biomed Data Sci. 2021;4:123–44.

258. Petch J, Di S, Nelson W. Opening the black box: the promise and limitations of explainable machine learning in cardiology. Can J Cardiol. 2022;38:204–13.

259. Jia Z, et al. The importance of resource awareness in artificial intelligence for healthcare. Nat Mach Intell. 2023;5:687–98.

260. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nat Med. 2022;28:31–8.

261. Torkamani A, Andersen KG, Steinhubl SR, Topol EJ. High-definition medicine. Cell. 2017;170:828–43.

262. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25:44–56.

263. Chen IY, et al. Ethical machine learning in healthcare. Annu Rev Biomed Data Sci. 2021;4:123–44.

264. Gao Y, Sharma T, Cui Y. Addressing the challenge of biomedical data inequality: An artificial intelligence perspective. Annu Rev Biomed Data Sci. 2023;6:153–71. https://doi.org/10.1146/annurev-biodatasci-020722-020704.

265. Newland JA. The All of Us Research Program: one size does not fit all. Nurse Pract. 2018;43:11.

266. Badal K, Lee CM, Esserman LJ. Guiding principles for the responsible development of artificial intelligence tools for healthcare. Commun Med. 2023;3:1–6.

267. Chen RJ, et al. Algorithmic fairness in artificial intelligence for medicine and healthcare. Nat Biomed Eng. 2023;7:719–42.

268. Norori N, Hu Q, Aellen FM, Faraci FD, Tzovara A. Addressing bias in big data and AI for health care: a call for open science. Patterns. 2021;2:100347.

269. Norgeot B, et al. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. Nat Med. 2020;26:1320–1324. Preprint at. https://doi.org/10.1038/s41591-020-1041-y.

270. Vasey B, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28:924–33.